HUMAN BIOLOGY
a record of research
FEBRUARY, 1950
VOL. 22 No. 1

RELATIONSHIPS BETWEEN AGE, VISUAL ACUITY AND COLOR VISION[1]

BY ALPHONSE CHAPANIS[2]

Department of Psychology,
The Johns Hopkins University,
Baltimore 18, Maryland

INTRODUCTION

THIS study is concerned with the relationships between age, visual acuity and color vision in an unselected sample of individuals. The first of these relationships, *viz.* that between age and visual acuity, has been investigated many times and so is of secondary interest to us here. The relationships between age and color vision, and between visual acuity and color vision, have not yet been determined adequately and so will receive primary emphasis in the present paper.

[1] This investigation was supported in part by funds supplied under Contract N5-ori-166, Task Order I, between the Special Devices Center, Office of Naval Research, and The Johns Hopkins University. This is Report No. 166-I-103, Project Designation No. NR-784-001, under that contract.

[2] The writer is indebted to many individuals who contributed greatly to this study. First and foremost among these are Mrs. Mary Lamb and Mr. Michael Leyzorek, both of whom supervised and did much of the testing reported here. In addition, Mrs. Lamb did the major portion of the statistical tabulations for this paper. Mr. Lynn G. Poole and Mr. G. V. Smith were primarily responsible

Previous studies on age and visual acuity

Although de Haan (9) appears to have first collected systematic data on the relationship between age and visual acuity, his study is overshadowed by Galton's extensive survey conducted about two decades later. de Haan studied 281 persons, ranging from 7 to 82 years of age, with Snellen's recently devised test chart. Starting each examinee at a great distance from the test chart, de Haan had the examinee walk forward slowly until he could just decipher the 20/20 letters. He found acuity to be greatest for the 10 to 20 year old individuals, somewhat less for the 20 to 30 year olds, and much less for people older than 30.

Galton made seventeen measurements, including visual acuity and "goodness of colour sense," on some 7,000 visitors to his first Anthropometric Laboratory at the Health Exhibition held at South Kensington in 1884. The methods of observation he used were described in 1885 (13), and the statistical analysis of his data was reported by Ruger and Stoessiger in 1927 (27). Briefly, his results show a marked curvilinear association between age and acuity: Visual acuity (the greatest distance at which diamond type could be read) is poor in young children, increases rapidly and reaches a maximum for individuals about 20 years old, declines very gradually from age 20 to about 45, and then more rapidly thereafter. With the exception of the data for the very young, this agrees closely with de Haan's findings. Essentially the same pattern has appeared in almost every study since then (see (2, 7, 19, 22, 23, 33 and 35), for example) and the more recent investigations—even though they have used more refined techniques—serve only to confirm the original observations made by de Haan and Galton.

Previous studies on age and color vision

Most studies on the relationship between age and color vision have been concerned with the development of color vision in infants and young children. A bibliography of researches on this problem is contained in an article by Smith (29), and two more recent articles on this relationship are those by Synolds and Pronko (34) and Pronko *et al.*

for the physical arrangements under which this study was conducted. Mrs. Marion Chapanis, Miss Grace Duvoisin, Mrs. Katherine P. Hill, Mr. Charles Barnes, Mr. Clarence H. Seaton, and Mr. Michael Wertheimer contributed generously of their time in administering the tests. Finally, Mr. Anthony de Lorenzo assisted in some of the later stages of the analysis.

(26). The gist of all these studies is that even infants are sensitive to color, but that children (up to the age of about 15) do not perform as well as adults on most color vision tests. A particular difficulty with the interpretation of these studies, and this point is well illustrated in the article by Pronko *et al.* (26), is that color vision capacity cannot be divorced from perceptual factors and test sophistication which greatly attenuate the performance of young children.

Studies of color vision capacity over wider age ranges are difficult to find. Galton's measure of "goodness of colour sense" is not particularly valid according to modern standards, and it appears that these data were never analyzed statistically. There are a few studies, (18) and (28) for example, which show that old people have more difficulty than younger people in distinguishing and recognizing blue-green, blue, and violet colors. Even though these are not rigorous quantitative studies, what they find is reasonable because the lens of the eye is known to become yellower with advancing years (10, 15). This yellowing of the lens has the effect of a filter which makes the discrimination of bluish colors somewhat more difficult and, in some cases, it may even lead to a form of color blindness resembling tritanopia, or blue-blindness (17). That this is probably a simple filtering effect, and not the result of deterioration in the retina itself, is suggested by Ferrara's data (11). He found that old people had higher thresholds for blue than younger people, but that old people with aphakic (lensless) eyes, had lower thresholds for blue than normal young people.

The present paper, however, is concerned with color vision tests which measure red-green discrimination and we shall turn our attention to studies on this function. Pierce (25) reports a correlation of —0.51 between age and performance on a sensitive color discrimination test—that is, old people do not do as well as young people. This figure, however, is based on the data from only 40 female color workers and the age range is not specified. Smith (29) studied 199 subjects, ranging from 5 to 87 years in age, with three color matching tests. His results show curvilinear functions very similar to that found by Galton for visual acuity. These data are open to serious question, however, because color-blind individuals seemed to do as well as, or better than, most color-normal individuals on the tests used. As a possible explanation for this anomalous result, one of the color-blind subjects who participated in the Smith study suggested to the author recently that he believes *his* matches were made entirely on the basis of the texture, and not the color, of the test samples.

Somewhat more recently, Tiffin and Kuhn (37) reported on an extensive survey of approximately 7,000 industrial employees. They used a four-item test of color vision and found a marked decrease in the percentage of individuals passing the test in the age span from 20 to 70 years. Although Tiffin and Kuhn did not identify or describe the test they used in this survey, it appears that the test consisted of cards manufactured by the Keystone Company and incorporated in their visual testing device, the Telebinocular. Virtually nothing is known of this device as a color vision test and there are several aspects of the findings which make the test highly suspect. Briefly, these are that (a) the validity and reliability of a test containing only four items are questionable (in this connection, see the data of Chapanis (4)), (b) some of the Tiffin-Kuhn data—in particular, the positive correlation between occupation and color vision score—suggest that this test may be measuring factors other than color vision, *e. g.,* intelligence, education, or, more generally, test sophistication, and (c) the high proportion of failures (26 per cent) for the best age group, 20-25 years, is completely at variance with results normally obtained with other color vision tests. In the discussion following the presentation of this paper before the Section of Ophthalmology of the American Medical Association, Dr. Tiffin himself stated that "we are not satisfied with this particular testing device for segregation of industrial employees."

Essentially these same objections to the four-plate Telebinocular test have been made by Boice, Tinker and Patterson (3). These investigators, in turn, studied the color vision capacities of 236 faculty men of all ranks at the University of Minnesota. For this study, they used the Ishihara and American Optical Company Tests and report that there was no evidence of deterioration in color vision capacity from age 20 to 59 years. They conclude with the statement that further research is indicated.

Previous studies on acuity and color vision

The relationship between visual acuity and color vision has received even less attention than the topics already reviewed. It is generally believed that refractive errors of large magnitude result in poor color discrimination, and there are many conjectures about this relationship in the literature (see (24) and (31), for example). However, as long ago as 1907, Köllner (21) reported that errors of refraction up to 20 diopters (with acuities corresponding to about 20/400) have no effect

on color mixtures although brightness is judged less well as the defect becomes more pronounced. Köllner's experiments, performed mainly on his own eyes, consisted in observing how accurately he could make a few color matches when he made himself artificially myopic by putting lenses of various powers in front of his eyes. When Barrett (1) tried essentially the same kind of experiment on his own eye he concluded that lenses which produced even a small decrease in form vision, to below 6/12 (20/40), make small colored lights unrecognizable. These, however, are qualitative observations with no systematic data to support them.

Of a different sort are some data of Collins (6) who reported that color blindness occurs less frequently in eyes which are apparently without demonstrable refractive error, and most frequently in eyes showing mixed astigmatism. Thus, he found 5.6 per cent of the "harmlessly color-blind" persons in his sample to have normal refraction, whereas none of the "dangerously color-blind" persons had normal refraction. Other comparisons of this sort are made for simple hyperopia, simple hyperopic astigmatism, compound hyperopic astigmatism, and so on. Collins' data can be easily misinterpreted, however, because they are, for the most part, expressed as percentages. In actuality, the numbers of cases involved are very small and none of the differences even approach statistical significance.

On the basis of extensive observations on "many" color normal individuals, Cronstedt (8) reports that persons with presbyopia, hyperopia and hyperopic astigmatism do not perform well on color vision tests unless they wear correction. "Mild" cases of myopia are not particularly hampered although the more severe cases of myopia have difficulty without their correction. Also important is the observation that these results vary for different tests: the Holmgren and Ishihara tests appear to be relatively immune to this effect whereas the Göthlin-Böstrom test is sensitive to it. Somewhat more recently, Chapanis (4) reported no relationship between age, visual acuity and results obtained on a number of color vision tests. The group of subjects he studied, however, was small (120 in number) and highly homogeneous with respect to both age and visual acuity.

Background of the present study

In December 1947, the author was invited to prepare an exhibit for the Baltimore Sesquicentennial Exhibition and this seemed like an

unusual opportunity to collect extensive data on the controversial points reviewed above. Accordingly, a small enclosed testing booth was arranged for the exhibit so that tests could be administered to all visitors who volunteered. It is the purpose of this article to report on the results of this survey.

SUBJECTS, APPARATUS AND PROCEDURE

Subjects

The subjects in this investigation were a cross section of typical exhibition visitors. Although they were not selected, their presence at

TABLE 1

Classification by age, sex, and color of the 574 individuals participating in this study

AGE	WHITE		COLORED		TOTAL
	Male	Female	Male	Female	
75-79	3				3
70-74	2	1			3
65-69	1	2			3
60-64	5	2			7
55-59	9	11			20
50-54	18	17		1	36
45-49	23	11	1	1	36
40-44	23	23		1	47
35-39	30	35		3	68
30-34	21	22		1	44
25-29	38	23	2	2	65
20-24	28	41		9	78
15-19	51	23	6	10	90
10-14	45	12	8		65
5-9	5	3		1	9
Total	302	226	17	29	574

an exhibition of this sort probably indicated that they are not representative of people in general. Speaking at a very descriptive level, the examinees appeared to represent a great variety of educational backgrounds, intelligences and occupations. Table 1 provides some descriptive statistics of a more quantitative sort about the 574 individuals participating in this survey. Although not determinate from this table (because

of the grouping), the subjects ranged from 7 to 77 years in age. The largest single group of subjects was composed of white males.

The testing situation

All the tests were administered in a booth, 6′ x 6′ x 8′ in size, erected in the Fifth Regiment Armory which housed the exhibition. The test booth was completely enclosed with dark curtains so that no illumination, other than that provided inside the test booth, reached the test materials. In addition to controlling the source of illumination on the tests, the heavy curtains afforded complete privacy for the examination. The test materials were placed on an easel directly in front of the seated subject. The examinee's chair and easel were rigidly separated. Illu-

FIG. 1. THE VISUAL ACUITY TEST USED IN THIS STUDY
(APPROXIMATELY FULL SIZE IN THIS ILLUSTRATION)

mination for the tests was provided by two Macbeth daylight lamps (Type ALS 30) which furnished 30 foot-candles of light of about 5,000° K at the test surface. The visual acuity test was administered at a distance of 32 inches. This was measured by a string attached to the easel on which the tests were placed. At the appropriate distance along the string was a knot which the examinee was required to hold against his temple during the visual acuity test.

The tests

The visual acuity test

This was a sharp, photographic reduction of a standard visual acuity test chart (see Fig. 1). At the test distance used, the top line corresponds to a Snellen acuity rating of 0.19 and the lines successively below it correspond to visual acuities of 0.38, 0.56, 0.77, 1.00, 1.25, 2.00, and 2.67. No examinee was able to read the last line. A point worth empha-

sizing here is that this test measures acuity for the same distance at which the color vision tests were read.

The color vision tests

Five pseudo-isochromatic tests were employed in these examinations:

1. An experimental, 6-plate pseudo-isochromatic test devised by the author. Only a brief description of this test is available in the literature (5).

2. C. G. Boström and I. Kugelberg: *Tabulae Pseudo-Isochromaticae: Tavlor för Färgsinnesprövning.* Printed by Kifa, Stockholm, 1944. Distributed by the Nordiska Bokhandeln, Stockholm. This is a foreign test which has been little used in this country. It is important to note, however, that this test is constructed according to different principles from those used in the Ishihara and AO tests which are commonly used in this country. The following plates from this test were used: 1, 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 16, 17, and 18.

3. S. Ishihara: *Tests for Colour-Blindness* (9th Edition). Printed by Kanehara and Company, Tokyo, 1940. Distributed by the C. H. Stoelting Company, Chicago, Illinois. This test has been widely used in this country and it has been shown repeatedly to be one of the most reliable and valid tests on the market. In the examinations here the subjects were asked to read all plates with the exception of numbers 18 through 21, and 26 through 32.

4. I. Dvorine: *Dvorine Color Discrimination Screening Test.* Printed privately in Baltimore, Maryland, 1947. Two forms of this test, each containing two plates, were used in these examinations.

5. *Pseudo-Isochromatic Plates for Testing Color Perception.* Engraved and printed by the Beck Engraving Company, Inc., New York, 1940. Distributed by the American Optical Company, Southbridge, Massachusetts. Thirty-nine plates of this test were used. The subjects were asked to read all but numbers 26, 29, 33, 34, 45 and 46.

Worth noting is the wide diversity of color vision tests represented here. As is well known, the Ishihara and American Optical Company tests contain a number of plates in common. (In the selections used here, 11 plates were common to the Ishihara and AO tests.) Aside from this duplication, however, the tests in general are quite different. Thus, for example, the Boström-Kugelberg and Dvorine plates are unique in their design, differing in basic principles from those employed in the construction of the Ishihara and AO tests. The author's test, in turn,

has some features differing from all the others. This wide selection of color vision tests enhances the generality of conclusions reached in this report.

Procedures

The examinations were all conducted individually. After the subject was seated in his chair, certain items of personal information were requested. Following this, the examinee was instructed to pick up the string which was attached to the test easel, to place the knot on the end of the string firmly against his temple adjacent to his eye and to hold the string taut. The test booklet was then opened to the visual acuity test and the examinee was asked to read all the letters he could. If the examinee normally wore glasses, he was instructed to wear them throughout the examination.

Following the visual acuity test, the examinee was given these instructions: " Now I will show you a series of colored plates on which you may or may not see numbers. Tell us exactly what you see. If you see a number, name it. If you don't see a number, tell us that too. Take all the time you want and if you are not sure, guess." All questions about the numbers on the plates were parried with the query, " What do you see? " or " How does it look to you? " The examinees were given absolutely no help in identifying any of the numbers on the plates. Whenever an examinee gave two or more responses to the same plate, he was asked to name the " best " or " clearest " one. Only these " best choice " responses were scored. A strictly literal scoring procedure was used. In order to be counted correct, the examinee's response had to agree exactly with the answers given in the test manuals. All the examiners used in this investigation were carefully rehearsed in these procedures before the actual tests began.

General comments about the test situation

The general environment was not optimal for conducting an experiment. The test booth was noisy and hot; the examinees tended to be a little flustered even though every effort was made to put them at ease. Outside, in the Armory, there were occasional band concerts, speeches and a generally high noise level emanating from the various exhibits. In other respects, however, the test conditions were good. Each examiner was carefully instructed in procedure, the visitors were examined individually and with curtains closed so that they were not embarrassed by their performance on the various tests. The illumination was excellent

and was probably much better controlled than is typical of most color vision testing situations. The test distance was also more carefully standardized than is characteristic of many color vision studies.

On the whole, the behavior of the examinees was extremely good. The exhibit had been publicized as a scientific experiment and most individuals—by far the greatest percentage of them—were extremely attentive and responsive to the examiners. Virtually none of them took the test lightly.

RESULTS

Table 2 exhibits the scores earned on the entire battery of color vision tests by the 574 individuals in this survey. The data are classi-

TABLE 2

Total number of plates read correctly on all color vision tests by the 574 subjects

SCORE	MALES			FEMALES			ALL CASES
	White	Negro	Total	White	Negro	Total	
81-85	116		116	83	3	86	202
76-80	80	6	86	74	7	81	167
71-75	33	4	37	37	10	47	84
66-70	21	4	25	13	5	18	43
61-65	2		2	9	3	12	14
56-60	3	2	5	2		2	7
51-55	2		2	1	1	2	4
46-50	1		1	2		2	3
41-45	1		1	1		1	2
36-40	2		2				2
31-35							
26-30	1		1				1
21-25	1		1	1		1	2
16-20	2		2				2
11-15	14		14	3		3	17
6-10	21	1	22				22
1-5	2		2				2
Total	302	17	319	226	29	255	574

fied according to sex and, crudely, according to color. It will be observed from Table 2 that the scores earned by the entire group fall into two distinct distributions. On the one hand, there are the 528 individuals who saw correctly 36 (actually 39) or more of the entire

group of 85 plates. An entirely separate and quite distinct distribution is made up of the 46 individuals who saw 30 (actually 27) or fewer plates correctly. According to almost any criterion, individuals in the latter group would be classified "color blind." Following a somewhat more rigorous classification, these 46 individuals are probably true dichromats. Approximately 13.2 per cent of all males, and 1.6 per cent of all females, fell into this classification. These are considerably higher proportions of color-defective individuals than are normally encountered in a randomly selected population. Many examinees remarked that they had long suspected color deficiency and had decided to take these tests for that reason. This is another respect, then, in which this sample deviates from a normal, unselected population. This finding, however, does not vitiate the form of the statistical analysis used here.

It is of incidental interest to note that there were considerably more color deficient men than color deficient women in this sample. This is a well established finding, but as already has been indicated, the absolute percentages in this sample cannot be trusted for comparative purposes. Also of incidental interest is the finding that the Negroes earned somewhat poorer scores than did the white people on this test battery. The number of Negro subjects in this sample, however, is far too small for any satisfactory conclusions.

The relationship between age and visual acuity

Alternate ways of scoring acuity

There are many ways in which the acuity of an individual could have been expressed and, of these various methods, two were selected for analysis. The first, referred to hereafter as the visual acuity score, was determined simply by counting the total number of letters read correctly on the acuity test (Fig. 1). Although this is an unconventional way of reporting acuity data, it does have the advantage of giving a good spread among individuals. In addition, the internal consistency of this score, evaluated by the split-half correlation used by psychologists for determining the reliability of tests (36), is very high—the uncorrected r equals $+0.900$.

A more conventional way of reporting acuity is to determine the last line of letters a person can read and then to attach some number to this line. There are two elements of arbitrariness in such a measure. The first relates to the decision about the last line the examinee could

read because, unfortunately, most people give scattered responses. They may read all the letters on the sixth line and two letters on the seventh, or may read all but one of the letters on the sixth line, and so on. We arbitrarily defined the last line an examinee could read as the last line for which he read one half or more of the letters.

The second element of arbitrariness concerns the number to be attached to the last line the examinee read. The lines could be simply numbered 1, 2, 3, and so on. Or each line might be identified by its Snellen rating, by the visual angle subtended by the critical details of the letters in each line, or by a logarithmic score computed by means of the following equation:

$$\text{AGO score} = 10 - \log_2 A$$

where A is the visual angle subtended by critical details of the letters. The AGO score was derived by a research group working with acuity measures (30) and is supposed to have some theoretical advantages over other measures.

These measures are functional transformations of each other and so should give roughly the same results. The Snellen rating was finally selected as the most appropriate alternate measure because this transformation yielded a distribution of acuity scores conforming very closely to the shape of the normal probability curve. For correlational purposes there is some advantage in working with normally distributed variates. The second best approximation to a normal curve was obtained with the visual acuity score. Line number and AGO score yielded markedly skewed distributions with about the same amount of skewness, and visual angle gave the poorest approximation to a normal curve since it exaggerated the skewness apparent in Fig. 3.

The two measures of acuity used in this study—visual acuity score and Snellen rating for the last line read correctly—give the closest approximations to normal distributions, they correlate fairly low with each other (+0.876), and they give the lowest average intercorrelations with the other measures if we ignore the average intercorrelation for visual angle. In test work, low intercorrelations among measures are desirable because this indicates that somewhat different aspects of performance are being sampled.

The relationship between age and visual acuity score

Fig. 2 exhibits the data on age and visual acuity score. Low scores indicate poor acuity, high scores good acuity. The linear relationship between age and visual acuity score, measured by the Pearson product

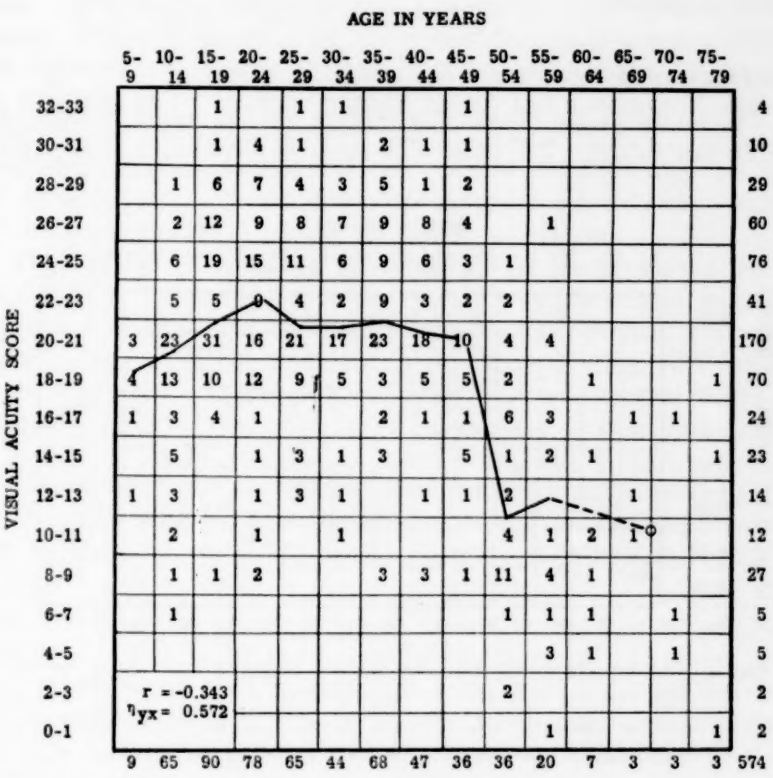

FIG. 2. THE RELATIONSHIP BETWEEN AGE AND VISUAL ACUITY SCORE (THE TOTAL NUMBER OF LETTERS READ CORRECTLY ON THE VISUAL ACUITY TEST SHOWN IN FIG. 1) FOR THE 574 SUBJECTS IN THIS SURVEY.

The regression line connects the medians of the columns. The last point is the median score for all persons over 60 years of age.

moment correlation, is —0.343. This means that old persons tend in general to earn somewhat poorer acuity scores than young persons. It is evident from an inspection of these data, however, that there is a marked curvilinear trend present there. Thus, subjects less than 15

years of age, and those older than 50, earned poorer scores than did those in the age range from 15 to 50. The curvilinear correlation η for the regression of acuity on age is 0.572. The value of 0.572 is significantly different from the linear r of —0.343 when tested statistically with the F test. We can be sure, therefore, that the curvilinear association is genuine.

The relationship between age and acuity

Fig. 3 exhibits the relationship between age and acuity measured by the Snellen rating. The linear r is equal to —0.231; η equals 0.440.

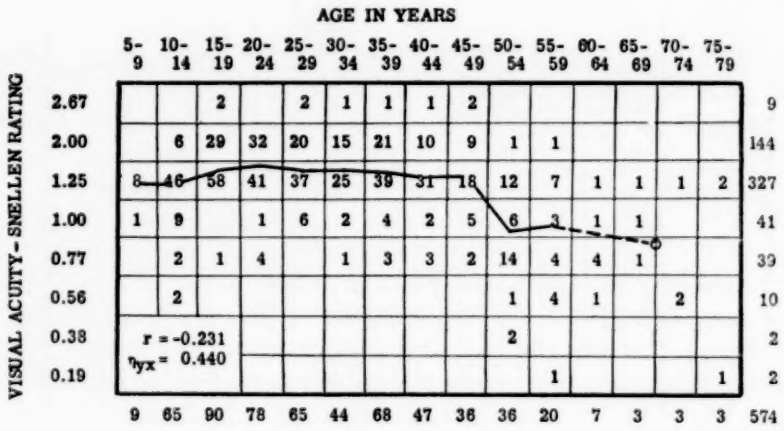

FIG. 3. THE RELATIONSHIP BETWEEN AGE AND SNELLEN RATINGS OF ACUITY FOR THE 574 SUBJECTS IN THIS SURVEY.

The regression line connects the medians of the columns. The last point is the median score for all persons over 60 years of age.

These values are quite different from the coefficients for age and visual acuity score reported above. There is still no question about the existence of a significant curvilinear relationship, but it is clear that the magnitude of the correlation depends markedly on how acuity is measured. It should be noted, incidentally, that these are the extreme values. The curvilinear correlation between age and the number of the last line read correctly is 0.550; age and visual angle, 0.525; and age and AGO score, 0.509.

The effect of heterogeneity on the correlations

The particular correlation values reported above are also dependent in part on the numbers of individuals falling in the various age groups. Thus, Galton's original sample of 6,992 people (27) had roughly the same composition as the sample studied here. His examinees ranged from 6 to 81 years as compared with a range of 7 to 77 years for this sample. The mean and standard deviation of the ages in Galton's sample were 29.9 and 12.6, respectively. In these data the corresponding values are 31.0 and 14.6. As in this study, Galton had a disproportionately small number of subjects younger than 11 or older than 60. It is not surprising, therefore, to find that the measures of relationship for this sample agree reasonably well with those computed for his data, —0.407 for the linear and 0.513 for the curvilinear correlation.

But that the distribution of individuals within the sample can markedly influence the correlation values is shown by comparable statistics computed for some data originating from Galton's Second Anthropometric Laboratory in the South Kensington Museum (20). These data are for 3,379 males measured in essentially the same way as were the 7,000 exhibition visitors. Although the age range for the two samples is identical (5 to 81 years), the second group was younger (average age = 24.3 years) and somewhat more homogeneous (standard deviation = 11.0) than the first. The correlations between age and visual acuity are also lower for the second sample: —0.311 for the linear r, 0.400 for η.

As was indicated earlier, however, the relationship between age and acuity is not one of the more important findings of this study. Its main importance is that it constitutes an internal check on the consistency and reliability of the measurements which were made in this study.

Summary

In summary, we can say that there is a marked curvilinear association between age and visual acuity. Very young and very old persons do not have as good acuity as those between the ages of 15 and 50. The magnitude of the curvilinear correlation is about 0.50 but it varies between 0.44 and 0.57 depending on how acuity is measured.

The relationship between age and color vision

Fig. 4 exhibits the relationship between age and the total number of color vision plates read correctly by the 574 individuals in this survey. The linear r for these data is equal to $+0.076$ and the curvilinear correla-

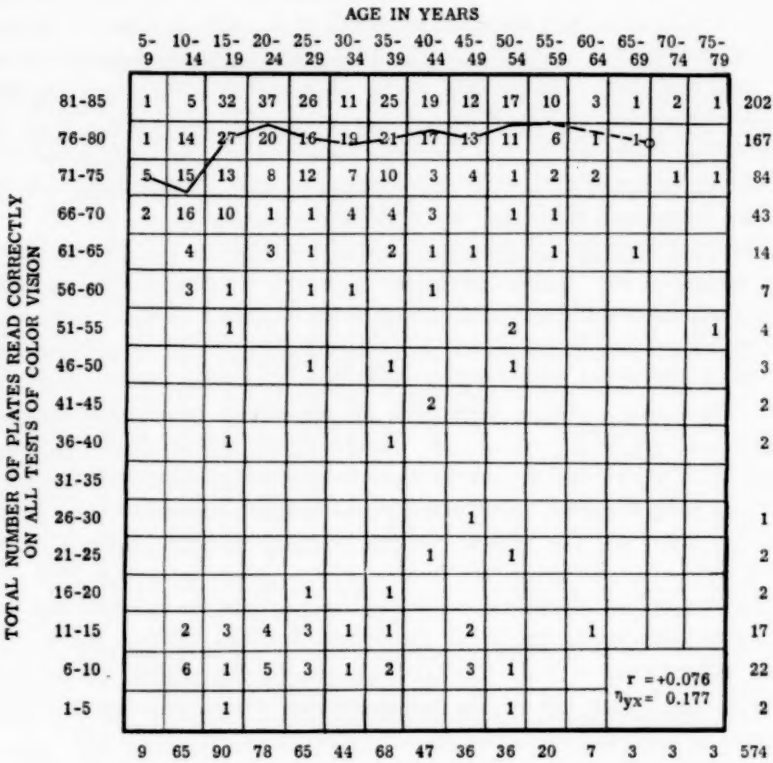

FIG. 4. THE RELATIONSHIP BETWEEN AGE AND THE TOTAL NUMBER OF PLATES READ CORRECTLY ON FIVE PSEUDO-ISOCHROMATIC TESTS OF COLOR VISION BY THE 574 SUBJECTS IN THIS SURVEY.

The regression line connects the medians of the columns. The last point is the median score for all persons over 60 years of age.

tion is equal to 0.177. Neither measure of correlation is significantly different from zero. This finding is of course suspect because of the very obvious bimodality in the color vision test data. For this reason, these data were divided at various arbitrary cut-off scores, replotted,

and correlations were then recomputed for the segregated data. The results of one such operation are exhibited in Fig. 5. This figure illustrates the relationship between age and color vision test score for all

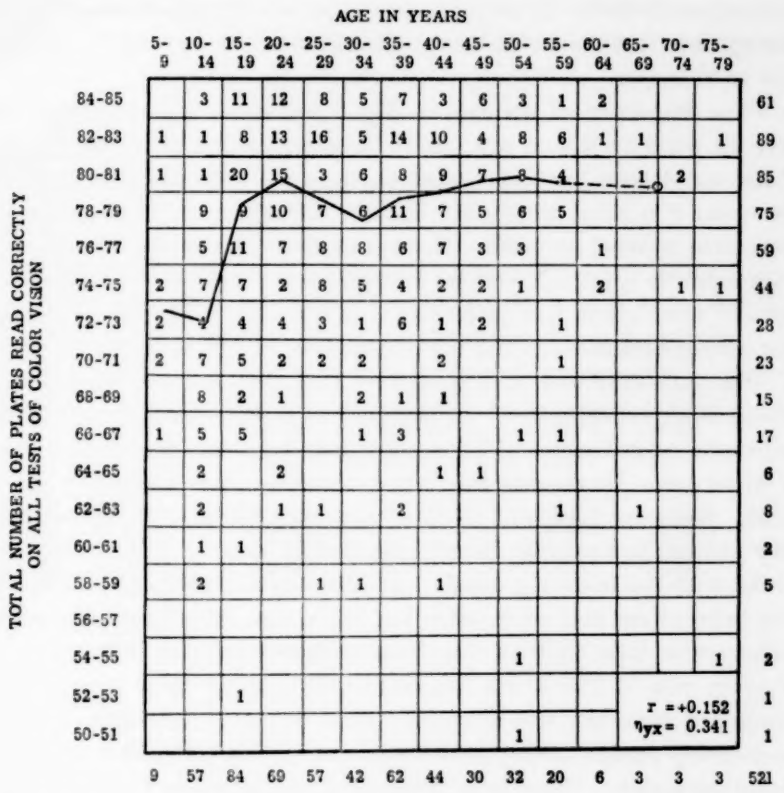

FIG. 5. THE RELATIONSHIP BETWEEN AGE AND THE TOTAL NUMBER OF PLATES READ CORRECTLY ON FIVE PSEUDO-ISOCHROMATIC TESTS OF COLOR VISION BY THE 521 SUBJECTS WHO READ 50 OR MORE PLATES CORRECTLY

This is an expanded plot of part of the data exhibited in Fig. 4. The regression line connects the medians of the columns. The last point is the median score for all persons over 60 years of age.

those individuals who earned color vision scores of 50 or more. This particular cut-off score was selected because it represents a break in the distribution of color vision scores which is not apparent from Fig. 4

because of the coarseness of the grouping employed there. For the data in Fig. 5, the correlation between age and color vision test score is still low. The linear r is equal to $+0.152$; η equals 0.341. When examined statistically with the F test, it appears that a curvilinear relationship is a significantly better fit to these data than is a linear relationship. It also appears that the curvilinear coefficient is significantly different from zero in this case.

Close inspection of the data in Fig. 5 will show that examinees who were 14 years or less in age turned in poorer scores than the average. If the correlations are recomputed omitting the data for these children, the linear r is equal to -0.001 and the curvilinear correlation for the same data is equal to 0.165. Neither of these is significantly different from zero. It appears, therefore, that whatever correlation exists between age and color vision test performance is due almost entirely to the fact that young examinees do not perform as well as adults on these tests.

The data exhibited in Fig. 4 were replotted and recomputed with several other cut-off scores. In one case, plots were made for all individuals who earned scores of 39 or more and for those who earned scores of 27 or less. In another case, the cut-off score was arbitrarily placed at 62. Since the results of all these recomputations came out the same, they will not be presented here in any further detail. They are all consistent with the following conclusion: There is a slight positive relationship between age and color vision test score, *i. e.*, in general, adults earn better scores than children do. This relationship drops almost to zero in every case if the scores for examinees younger than 15 years are eliminated from the computations. This indicates, therefore, that the relationship is due almost entirely to the poorer scores earned by the young examinees. There is no evidence that test scores decrease for the old examinees.

Correlations between age and color vision scores for individual tests

Because there is some likelihood that individual color vision tests might give different results, correlations were computed between age and scores on single tests with the results shown in Table 3. The first two columns give the linear and curvilinear correlations between age and scores on single tests for all 574 cases. Only one of them is significantly different from zero.

The scores on the individual color vision tests were all bimodal in essentially the same manner as the scores for all tests shown in Fig. 4. For each test, therefore, the lowest group of scores was eliminated and the correlations were recomputed with markedly skewed, but unimodal, distributions similar to that shown in Fig. 5. The cut-off scores were set at the lowest value which would eliminate the bimodality; they were

TABLE 3

Linear (r) and curvilinear (η) correlations between age and scores on individual tests of color vision

The correlations in the first two columns are for the entire group of 574 subjects; those in the last two columns are for the roughly 530 cases remaining after the very low scores are excluded (see the text for a fuller explanation).

TEST	ALL CASES		MINUS LOW SCORES	
	r	η	r	η
AC	+0.005	0.144	—0.070	0.202
AO	+0.101**	0.201	+0.098	0.293*
BK	+0.039	0.156	—0.017	0.190
Dvorine	+0.021	0.149	—0.030	0.148
Ishihara	+0.085	0.165	+0.151*	0.215
All tests	+0.076	0.177	+0.100**	0.264*

* Significant at the 1% level

** Significant at between the 5% and 1% levels

not selected to eliminate all color deviant individuals from the correlations. For the author's test (AC test), all zero scores were excluded; for the AO test, scores of 15 or less; for the BK test, scores of 0 and 1; for the Dvorine test, scores of 0; and for the Ishihara, scores of 8 or less were excluded. The cut-off score for all tests was at 38. The new correlation values are shown in the last two columns of Table 3.

The coefficients in Table 3 are all uniformly small and, with few exceptions, not significantly different from zero. We must conclude, therefore, that the analysis by individual tests does not alter our conclusions on the association between age and color vision.

The relationship between visual acuity and color vision test performance

Fig. 6 exhibits the relationship between visual acuity score and the total number of color vision plates read correctly on all tests. The

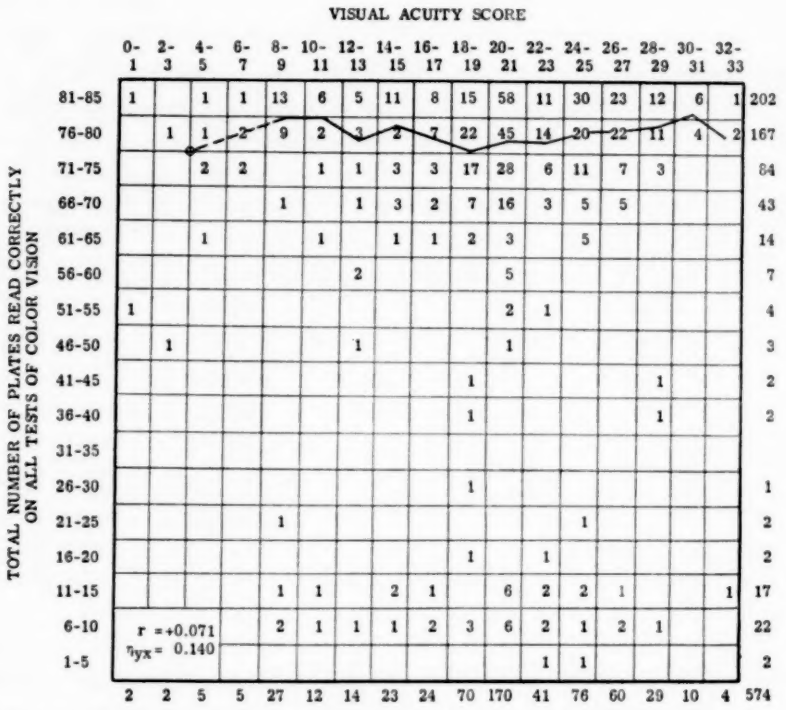

FIG. 6. THE RELATIONSHIP BETWEEN VISUAL ACUITY SCORE AND THE TOTAL NUMBER OF PLATES READ CORRECTLY ON FIVE PSEUDO-ISOCHROMATIC TESTS OF COLOR VISION BY THE 574 SUBJECTS IN THIS SURVEY

The regression line connects the medians of the columns. The first point is the median color vision score for all persons who had acuity scores of 7 or less.

linear r for this set of data is equal to $+0.071$; the curvilinear relationship is expressed by the coefficient 0.140. Neither of these is significantly different from zero. As before, the data were regrouped in various ways

and Fig. 7 exhibits the results of one such regrouping. This figure shows the relationship between visual acuity scores and color vision test scores for the 528 individuals who earned color vision scores of 39 or greater. The linear r for these data is $+0.068$; the curvilinear relationship is

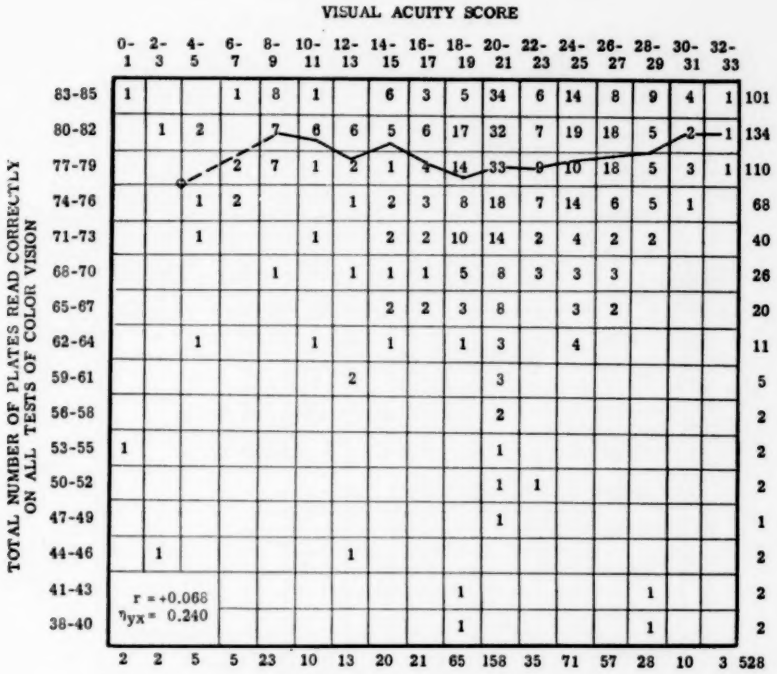

FIG. 7. THE RELATIONSHIP BETWEEN VISUAL ACUITY SCORE AND THE TOTAL NUMBER OF PLATES READ CORRECTLY ON FIVE PSEUDO-ISOCHROMATIC TESTS OF COLOR VISION BY THE 528 SUBJECTS WHO READ 39 OR MORE PLATES CORRECTLY

This is an expanded plot of part of the data exhibited in Fig. 6. The regression line connects the medians of the columns. The first point is the median color vision score for all persons who had acuity scores of 7 or less.

equal to 0.240. The latter correlation is significantly different from zero. As in the case of the data for the relationship between age and color vision test scores, these data were replotted and recomputed for cut-off scores at 50 and at 62. None of these recomputations yielded a significant relationship between visual acuity score and color vision test score.

The relationship between Snellen ratings of acuity and color vision scores

The relationship between Snellen ratings of acuity and color vision scores is exhibited in Fig. 8 for the entire group of 574 cases. Figure 9 shows the data for the 528 cases who saw 39 or more plates correctly

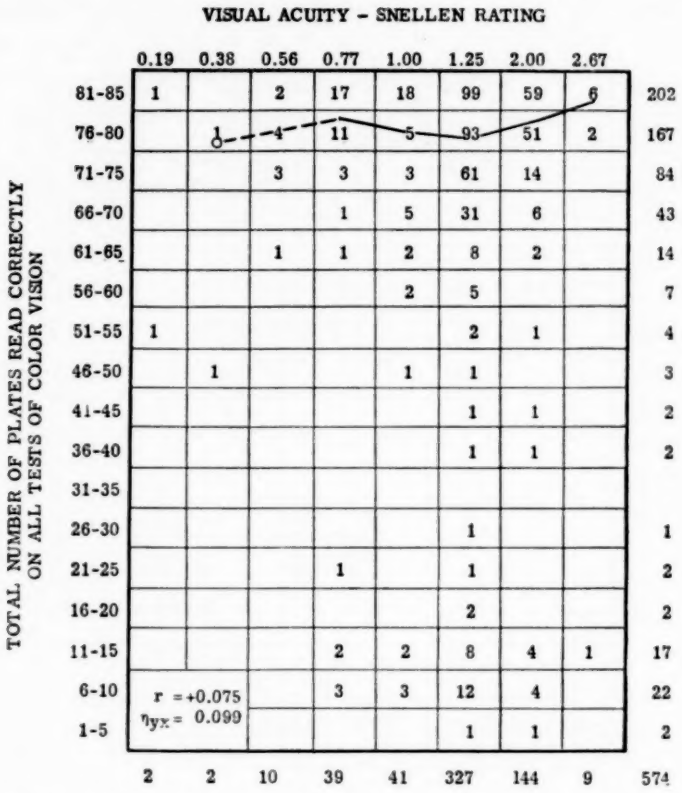

FIG. 8. THE RELATIONSHIP BETWEEN SNELLEN RATINGS OF ACUITY AND THE TOTAL NUMBER OF PLATES READ CORRECTLY ON FIVE PSEUDO-ISOCHROMATIC TESTS OF COLOR VISION BY THE 574 SUBJECTS IN THIS SURVEY

The regression line connects the medians of the columns. The first point is the median color vision score for all persons who had Snellen acuity ratings of 0.56 or less.

on all color vision tests. Neither correlation coefficient is significantly different from zero in the former instance; both are significant at the 1 per cent level in the latter instance. If one excludes the markedly color deviant individuals, therefore, there does seem to be some positive

association between visual acuity and performance on these tests: People with better acuity tend to make somewhat fewer errors.

But the demonstration of this statistically significant relationship is not, in the writer's opinion, the most important finding. Study of

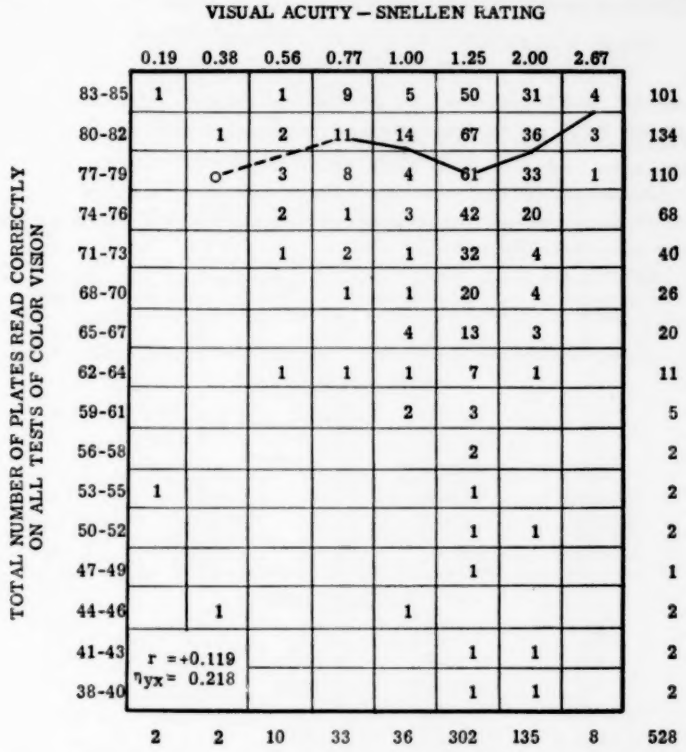

FIG. 9. THE RELATIONSHIP BETWEEN SNELLEN RATINGS OF ACUITY AND THE TOTAL NUMBER OF PLATES READ CORRECTLY ON FIVE PSEUDO-ISOCHROMATIC TESTS OF COLOR VISION BY THE 528 SUBJECTS WHO READ 39 OR MORE PLATES CORRECTLY

This is an expanded plot of part of the data exhibited in Fig. 8. The regression line connects the medians of the columns. The first point is the median color vision score for all persons who had Snellen acuity ratings of 0.56 or less.

the correlation measures, and of the raw data and regression lines in Figs. 8 and 9, cannot fail to impress the reader with how poor the association is. Indeed, we might well conclude that the relationship between acuity and color vision is *statistically* significant, but so low as to be *practically* insignificant. Note that two of the four persons with

the poorest acuity read many more color vision plates correctly than did most people with considerably better acuity.

It is worth noting in passing that in these, and in the correlations given in Table 4, visual acuity is considered the independent variate. Curvilinear correlations are computed only for the regression of color

TABLE 4

Linear (r) and curvilinear (η) correlations between visual acuity scores (upper half), or Snellen ratings of acuity (lower half), and scores on individual tests of color vision

The correlations in the first two columns are for the entire group of 574 subjects; those in the last two columns are for the roughly 530 cases remaining after very low scores are excluded.

TEST	ALL CASES		MINUS LOW SCORES	
	r	η	r	η
	Visual acuity scores			
AC	+0.083	0.137	+0.117*	0.239*
AO	+0.081	0.151	+0.096**	0.222**
BK	+0.097**	0.159	+0.125*	0.257*
Dvorine	+0.053	0.147	+0.012	0.166
Ishihara	+0.019	0.136	—0.071	0.226**
All tests	+0.071	0.140	+0.068	0.240**
	Snellen ratings			
AC	+0.060	0.086	+0.076	0.187*
AO	+0.091	0.113	+0.152*	0.209*
BK	+0.096**	0.122	+0.143*	0.197*
Dvorine	+0.046	0.111	+0.044	0.122
Ishihara	+0.025	0.084	—0.019	0.186**
All tests	+0.075	0.099	+0.119*	0.218*

* Significant at the 1% level

** Significant at between the 5% and 1% levels

vision score on acuity. This means, further, that the linear correlations between acuity and color vision scores could not exceed the curvilinear coefficients even if any of the alternate measures of acuity discussed earlier had been used.

Relationship between acuity and color vision scores for individual tests

Correlation coefficients between visual acuity scores, or Snellen ratings of acuity, and scores on individual tests of color vision are listed in

Table 4. As in Table 3, one-half of these coefficients were computed for the roughly 530 individuals remaining after the lowest scores were excluded to eliminate the bimodality of these distributions. Roughly one-third of the coefficients in this table are significant at the 5 per cent level or better and we must conclude that there is some evidence here for a positive association between acuity and performance on color vision

TABLE 5

Tabulation of visual acuity scores for examinees who do not wear glasses, those who wear glasses and wore them for these tests, and those who wear glasses but did not have them to wear for these tests

VISUAL ACUITY SCORE	DO NOT WEAR GLASSES		WEAR GLASSES AND Wore them during these tests		WEAR GLASSES AND Did not wear them during these tests	
	N	%	*N*	%	*N*	%
32-33	1	0.3	2	0.9	1	2.0
30-31	9	3.1	1	0.4		
28-29	15	5.2	11	4.8	2	4.1
26-27	41	14.1	16	7.0	3	6.1
24-25	48	16.6	23	10.1	5	10.2
22-23	25	8.6	11	4.8	5	10.2
20-21	87	30.0	69	30.3	11	22.4
18-19	40	13.8	20	8.8	10	20.4
16-17	7	2.4	14	6.1	2	4.1
14-15	10	3.4	9	3.9	4	8.2
12-13	3	1.0	9	3.9	2	4.1
10-11	3	1.0	9	3.9		
8-9	1	0.3	22	9.6	2	4.1
6-7			4	1.8	1	2.0
4-5			5	2.2		
2-3			2	0.9		
0-1			1	0.4	1	2.0
Total	290	99.8	228	99.8	49	99.9

tests. As before, however, we must also conclude that the coefficients are so small as to be practically insignificant. There do not seem to be outstanding differences among the various tests in the amount of this relationship.

Relationship between refractive error and color vision score

As a further check on the possibility of a relationship between refractive error and color vision test performance, the 574 individuals

in this survey were separated into three groups: (1) those individuals who do not wear glasses; (2) those who wear glasses habitually and wore them for this examination; and (3) those who wear glasses but did not wear them for this examination. The visual acuity scores for these three groups of individuals are exhibited in Table 5. As might be expected, those examinees who do not wear glasses earned considerably better visual acuity scores than did those individuals who normally wear

TABLE 6

Total number of color vision plates read correctly by the same examinees as are listed in Table 5

COLOR VISION SCORE	DO NOT WEAR GLASSES		WEAR GLASSES AND			
			Wore them during these tests		Did not wear them for these tests	
	N	%	*N*	%	*N*	%
81-85	84	29.0	94	41.2	19	38.8
76-80	84	29.0	64	28.1	18	36.7
71-75	54	18.6	23	10.1	6	12.2
66-70	31	10.7	11	4.8	1	2.0
61-65	7	2.4	6	2.6	1	2.0
56-60	4	1.4	2	0.9	1	2.0
51-55	1	0.3	3	1.3		
46-50	1	0.3	2	0.9		
41-45	1	0.3	1	0.4		
36-40	1	0.3	1	0.4		
31-35						
26-30	1	0.3				
21-25			2	0.9		
16-20	1	0.3	1	0.4		
11-15	9	3.1	7	3.1	1	2.0
6-10	9	3.1	11	4.8	2	4.1
1-5	2	0.7				
Total	290	99.8	228	99.9	49	99.8

glasses and wore them for this examination. (Parenthetically, of those persons who do not wear glasses, one had a Snellen rating of 0.56, 3 had ratings of 0.77, and the remaining 286 had ratings of 1.00 or better.) Many examinees who wear glasses, however, earned visual acuity scores which equalled those of the people who do not wear glasses. This also is to be expected if the correction is adequate for the individual concerned. It should be noted also that this information was inadvertently omitted from the data sheets for seven individuals.

Table 6 exhibits the color vision test scores earned by the three

classes of individuals. There is so much overlap between the color vision test scores earned by the three groups that it would be extremely hazardous to venture any general statement about differences in color vision capacity on the basis of this tabulation. Of the 290 persons who do not wear glasses, 7.5 per cent were clearly color deficient (saw fewer than 30 plates correctly); of the 277 persons who wear glasses, 8.7 per cent were clearly color deficient. These percentages do not differ significantly. There is, therefore, no obvious relationship between color vision test score and refractive error as indicated by the wearing of corrective eye glasses.

DISCUSSION

The data on the relationship between age and visual acuity are about what we should have expected on the basis of earlier studies and so merit little discussion. A point of some incidental interest, however, is the finding that the actual size of the correlation coefficients can be markedly altered depending on how acuity is expressed. This does not seem to have been pointed out in any earlier studies and it should at least caution us against an implicit faith in the absolute magnitudes of such coefficients. The almost complete lack of association between age and color vision scores (and between acuity and color vision scores) can be interpreted in at least four ways: (a) there really is no (or at best a very low) correlation between these measures, (b) there really is a relationship but our sample was not representative and so obscured the correlation, (c) there really is a relationship but we did not find it because we used the wrong kinds of tests, or (d) there really is a relationship but the form of the statistical analysis obscured it. Although all the evidence in this study supports alternative (a), we must say a few words about the other possibilities.

Representativeness of the sample

The generality of the conclusions reached in a study such as this one depends greatly on the representativeness of the sampling. Evidence has been presented in the earlier sections of this article to indicate that the individuals in this survey are not a random sample of people in general. On the other hand, the sample was extremely heterogeneous in composition. It is difficult to escape this problem (Galton himself faced it) and perhaps the best assurance of generality is that of agreement with other studies, conducted with other samples of people and other techniques. Thus, the curvilinear relationship between age and

visual acuity represents an internal check on the reliability of the data reported here. This kind of relationship has been discovered in many other studies and it is likely, therefore, that the sample of individuals in this experiment is representative at least to that extent.

Although the individuals in this survey were obviously heterogeneous with respect to age and color vision capacity, it might appear offhand that they all had remarkably good visual acuities. Thus, the two subjects who had the poorest acuity were just barely able to decipher the largest letter on the test and so received Snellen ratings of 0.19. This is not as bad as one is apt to discover in a random population. Karpinos (19), for example, reports that about 8 per cent of World War II selectees (ranging in age from 18 to 39) had visual acuity ratings worse than 0.20 on a test of far acuity. The situation is a little more complex than it might appear, however, because Snellen ratings of acuity are quite arbitrary—even though they are widely accepted and used—and they can be shown to vary enormously with a variety of external conditions. Illumination is one such important factor. An individual who receives a rating of 1.00 at one time may easily receive a rating of 0.50 or 1.50 if the illumination is varied.

It is apparent from Fig. 3 that nearly 85 per cent of the 574 examinees received acuity ratings better than 1.00—the hypothetical norm. This is probably due to the high level of illumination provided for these tests. It means that these ratings must be revised downwards if they are to correspond with the ratings in a random population. Although there is no simple conversion factor to apply here for this purpose, this discussion should at least make it clear that the visual acuities constitute a more representative sample than one might at first suppose. Nonetheless, we should exercise some caution in interpreting the acuity relationships. This sample does not include any individuals with extremely high refractive errors, such as an ophthalmologist might see in his practice. It is possible that patients with extremely gross refractive defects may have difficulty with tests of color vision such as were used here.

Adequacy of the tests used

The third alternative explanation for the low correlations reported here also merits some discussion. Although we have described one of the tests as a "visual acuity" test, we do not mean to imply that this is a test of pure retinal resolution. At such a close distance (32 inches), the test scores are undoubtedly weighted heavily by the factor of accom-

modation which is known to decrease markedly in old age (10, 12). On the other hand, a test of near visual acuity was purposely selected precisely because near and far acuity do not correlate very highly (14, 32). This study has shown only that variations in form vision—whatever the source of these variations—are largely independent of variations in scores on color vision tests when both kinds of tests are administered at the same distance.

There are some objections that one can make to the color vision test scores used in this investigation. It has long been known that standard pseudo-isochromatic tests of color vision were designed primarily as dichotomous tests. That they largely function in this manner is evident from the bimodal distributions of color vision test scores as exhibited in Table 2 and Fig. 4. Even so, scores on these tests are very roughly quantitative, as has been shown by Hardy *et al.* (16) and Wright (38), but they can hardly be said to differentiate various amounts of color defect with equal exactitude.

In addition, it is known that these pseudo-isochromatic tests do not test pure color vision capacity. They are also perceptual tests to a fairly large extent (see Chapanis (4), for example). On the other hand, they are the best single measures of color vision capacity available to the visual examiner at the present time. It is also worth repeating that this investigation employed a considerably better and larger selection of pseudo-isochromatic tests than has ever been used before. In view of this last consideration, it is indeed a most significant finding that the correlations between color vision test scores and age, and between color vision scores and visual acuity, are essentially zero. In the author's opinion, however, these zero correlations do not show that such relationships do not exist, but rather that pseudo-isochromatic tests are perhaps too crude to elicit them. It is most likely that further work on this problem will be profitable only when more sensitive tests of color vision have been devised. One thing is clear, however: Errors made by adults on present pseudo-isochromatic tests of color vision are largely independent of age and minor defects in visual acuity.

The form of the statistical analysis

These are difficult data to summarize statistically. Most of the distributions are bimodal and even when bimodality has been eliminated the distributions are markedly skewed. Correlational statistics have been used because they are convenient, but their validity for data like these is open to some question. The author has assumed throughout

that a low r and low η must mean that the regression of Y on X is essentially flat and parallel to the X-axis. In any event, Figs. 2 through 9 provide the raw data and so guard against the possibility that important relationships may have been obscured by the form of the statistical analysis.

Some observations with artificial myopia

At first glance, it appears that there "ought to be" some relationship between acuity and color vision test performance. As a check on the validity of the findings reported here, I have made some observations with the technique used by Köllner and Barrett. The surprising result is that artificial myopia really does not affect one's ability to read pseudo-isochromatic plates. By introducing lenses before my eye, I can make myself so myopic that I cannot read a 20/100 acuity letter, but this does not impair my ability to read the color vision plates.

The explanation for this is apparent as soon as one tries the experiment. Pseudo-isochromatic plates are composed of groups of colored dots and the task for the examinee is to integrate, or perceive, the dots of one color as a number against a background of dots of some other color or colors. A little myopia blurs the outlines of the dots so that for some plates it is actually easier to fuse them into a continuous pattern. Of course, with enough lens power in front of one's eye, it is impossible to see anything at all. These observations have been confirmed by two other subjects and they lend support to the findings of this paper that slight defects in form vision do not impair ability to read pseudo-isochromatic plates. It is still entirely possible that lantern tests which use point sources of light, as Barrett did in his study, might be sensitive to defects in acuity.

SUMMARY AND CONCLUSIONS

Five hundred and seventy-four visitors to the Baltimore Sesquicentennial Exhibition were tested with a near visual acuity test chart and with five pseudo-isochromatic tests of color vision. The visitors ranged from 7 to 77 years of age. The results of this investigation show that:

1. There is a marked curvilinear relationship between age and visual acuity. Very young and very old examinees do not attain as good visual acuity ratings as do those individuals in the age range of 15 to 50.

2. There is a slightly positive relationship between age and color vision test score. This relationship appears to be due entirely to the

fact that very young examinees do not earn as high scores as adults do. When recomputed for individuals 15 years and older, the correlations are all zero. There is no evidence that very old examinees are handicapped on these color vision tests.

3. There is some slight evidence of a positive relationship between visual acuity and color vision test scores. Although several statistically significant correlations were found, the correlations are all so small as to be practically insignificant. It is concluded that errors on present pseudo-isochromatic tests of color vision are largely independent of minor defects in form vision.

REFERENCES

1. Barrett, J. W. The problem of the visual requirements of the sailor and the railway employee. *Rep. Brit. Assoc. Adv. Sci.*, 1915, *84*, 256-263.
2. Boerma, D. and K. Walther. Untersuchungen über die Abnahme der Sehschärfe im Alter. *Archiv f. Ophthal.*, 1893, *39*, 71-82.
3. Boice, M. L., M. A. Tinker, and D. G. Patterson. Color vision and age. *Amer. J. Psychol.*, 1948, *61*, 520-526.
4. Chapanis, A. A comparative study of five tests of color vision. *J. Opt. Soc. Amer.*, 1948, *38*, 626-649.
5. ———. An attempt to construct a quantitative pseudo-isochromatic test of color vision. *Amer. Psychologist*, 1948, *3*, 245.
6. Collins, G. L. Color Blindness: Its Relation to Other Ocular Conditions, and the Bearing on Public Health of Tests for Color Sense Acuity. *Publ. Hlth. Bull. No. 92. Washington*: Gov't Printing Office, 1918. 29 pp.
7. Collins, S. D. and R. H. Britten. Variation in eyesight at different ages, as determined by the Snellen test. *Publ. Hlth. Rep.*, 1924, *39*, 3189-3194.
8. Cronstedt, L. Några iakttagelser angående behovet av glaskorrektion vid färgsinnesundersökningar. *Svenska Läkartidn.*, 1937, *34*, 145-150.
9. de Haan, V. Onderzoekingen naar den invloed van der leeftijd op de gezigtsscherpte. Diss. inaug. Utrecht, 1862.
10. Donders, F. C. On the Anomalies of Accommodation and Refraction of the Eye. *London*: The New Sydenham Society, 1864.
11. Ferrara, A. Soglia visiva per il bleu nei giovani, nei vecchi e negli afachici. *Ann. Ottalmol. e Clin. Oculist.*, 1940, *68*, 147-154.
12. Friedenwald, J. S. The eye. Chapter 20, pp. 535-555, in Cowdry, E. V. (Ed.), Problems of ageing (2nd Ed.). *Baltimore*: Williams and Wilkins, 1942.
13. Galton, F. On the Anthropometric Laboratory at the late International Health Exhibition. *J. Anthrop. Inst.*, 1885, *14*, 205-221.

14. GIESE, W. J. The interrelationship of visual acuity at different distances. *J. Appl. Psychol.*, 1946, *30*, 91-106.

15. HALLAUER, O. Ueber die Absorption von Kurtzwelligem Licht durch die menschliche Linse. *Klin. Monatsbl. f. Augenhk.*, 1909, *47*, 721-730.

16. HARDY, LEG. H., G. RAND, and M. C. RITTLER. The Ishihara test as a means of detecting and analyzing defective color vision. *J. Gen. Psychol.*, 1947, *36*, 79-106.

17. HESS, C. Ueber "Blaublindheit" durch Gelbfärbung der Linse. *Archiv. f. Augenhk.*, 1908, *61*, 29-37.

18. ———. Messende Untersuchungen über die Gelbfärbung der menschlichen Linse und über ihren Einfluss auf das Sehen. *Archiv f. Augenhk.*, 1909, *63*, 164-180.

19. KARPINOS, B. D. Visual acuity of selectees and army inductees. *Human Biol.*, 1944, *16*, 1-14.

20. KOGA, Y., and G. M. MORANT. On the degree of association between reaction times in the case of different senses. *Biometrika*, 1923, *15*, 346-372.

21. KÖLLNER, H. Über den Einfluss der Refractionsanomalien auf die Farbenwahrnehmung, besonders auf die Beurteilung spectraler Gleichungen. *Ztschr. f. Augenhk.*, 1907, *18*, 430-441.

22. MILES, W. R. Abilities of older men. *Person. J.*, 1933, *11*, 352-357.

23. ———. Training, practice and mental longevity. *Science*, 1935, *81*, 79-87.

24. MURRAY, E. Color vision tests. Pp. 275-282 in GLASSER, O. (Ed.), Medical physics. *Chicago*: Year Book Publishers, 1944.

25. PIERCE, W. O'D. The Selection of Colour Workers. *London*: Sir Isaac Pitman and Sons, Ltd., 1934.

26. PRONKO, N. H., J. W. BOWLES, JR., F. W. SNYDER, and D. L. SYNOLDS. An experiment in pursuit of "color-blindness." *J. Genet. Psychol.*, 1949, *74*, 125-142.

27. RUGER, H. A. and B. STOESSIGER. On the growth curves of certain characters in man (males). *Ann. Eugen.*, 1927, *2*, 76-110.

28. RYDIN, H. De la capacité de distinguer, aux différents ages de la vie, les radiations vertes ayant une longeur d'onde de 539$\mu\mu$ et celles vert-bleuatre de 500$\mu\mu$. *Comptes rendus Soc. Biol.*, 1927, *96*, 818-820.

29. SMITH, H. C. Age differences in color discrimination. *J. Gen. Psychol.*, 1943, *29*, 191-226.

30. STAFF, PERSONNEL RESEARCH SECTION, THE ADJUTANT GENERAL'S OFFICE, DEPARTMENT OF THE ARMY. Studies in Visual Acuity. *PRS Report No. 742*. *Washington*: Government Printing Office; 1948.

31. STEINMARK, H. Future problems of research in colour vision. *Dioptric Rev. and Brit. J. Physiol. Opt.*, 1942, *3*, 114-125.

32. SULZMAN, J. H., E. B. COOK, and N. R. BARTLETT. The reliability of visual acuity scores yielded by three commercial devices. *J. Appl. Psychol.*, 1947, *31*, 236-240.

33. SYDENSTRICKER, E., and R. H. BRITTEN. The physical impairments of adult life. Prevalence at different ages, based on medical examinations by the Life Extension Institute of 100,924 white male life insurance holders since 1921. *Amer. J. Hyg.*, 1930, *11*, 95-135.

34. SYNOLDS, D. L., and N. H. PRONKO. An exploratory study of color discrimination of children. *J. Genet. Psychol.*, 1949, *74*, 17-21.

35. THOMPSON, L. R., and R. H. BRITTEN. Standards of measurements of ten thousand male workers: preliminary note, with special reference to racial factors. *Amer. J. Publ. Hlth.*, 1924, *14*, 383-390.

36. THORNDIKE, R. L. Personnel Selection: Test and Measurement Techniques. *New York*: Wiley, 1949.

37. TIFFIN, J., and H. S. KUHN. Color discrimination in industry. *Arch. Ophthal.*, 1942, *28*, 851-859.

38. WRIGHT, W. D. Researches on Normal and Defective Colour Vision. *St. Louis*: Mosby, 1947.

THE PHYLOGENY OF DIGITAL PATTERNS IN MAN AND ITS BEARING ON RACIAL AFFINITIES: A STUDY IN HUMAN ECOLOGY

BY ROY I. WOLFE
Department of Geography
University of Toronto

INTRODUCTION

DERMATOGLYPHICS have been offered as criteria in the racial[1] classification of man. Present doubts as to their value for such a purpose stem in large part from difficulties raised by existing assumptions about their phylogeny. In the present paper the assumptions for one set of dermatoglyphics, the digital patterns, are reviewed, and alternative ones suggested, in the hope that these difficulties may be resolved.

Boyd (1940, pp. 334-335) lists a number of "requirements for a valuable criterion of classification," and Cummins and Midlo (1943, pp. 254-255) discuss the extent to which dermatoglyphics satisfy them. Their discussion follows:

> "(a) The criteria must be objective, so that the individuals will be classified consistently by different observers. The methodology of dermatoglyphics is now so standardized that constancy of determinations is sufficiently assured. As a result of intensive interest in finger prints, descriptive methods for these features progressed rapidly to the point of stability, and abundant data from different races are on record. In the case of

[1] It is well for anyone who uses the term "racial" to explain what he means by it. By my definition (Wolfe, 1947, p. 3), "racial distinctions exist between two human populations when there are stable and demonstrable phenotypical differences between their norms, these differences having been made possible by the long-continued geographical isolation from each other of these two populations down to present or recent (pre-Columbian) times, so that different lines of development have been followed and different mutations have arisen and been established."

palm and soles less material has been accumulated; additionally, some of the early records, based upon methods now superseded, admit at best only tentative translation into terms of current methods.

(b) The traits used in racial classification must not be subject to extensive environmental modification. This requirement is met by dermatoglyphics, which are not modifiable post-natally.

(c) Ideally a criterion chosen for racial comparison should be controlled by a simple and known genetic process. Dermatoglyphic features, unfortunately, are like most other traits which figure in racial classifications, being regulated by a complex and incompletely understood genetic mechanism.

(d) The traits should be non-adaptive, thus eliminating effects of natural selection. Dermatoglyphics satisfy this condition probably as well as any physical characteristic, and better than many. The superiority of patterns over patternless configurations, both in heightening frictional resistance and in increasing tactile acuity is recognized . . ., but it is extremely doubtful that the difference between a pattern and a patternless configuration or between two pattern types has any selective value whatsoever. Grüneberg . . . concedes the improbability of a direct selection of qualities of dermatoglyphics, but he holds that selection may still have been indirectly responsible for effecting racial differences in these features. He asserts, without offering evidence, that there may be genetic linkage with traits which do have a selective value. Actually, dermatoglyphics are not thus correlated with characteristics that have been tested . . ., and their genetic independence (from head form, nasal form, general body build and blood groups) adds a distinct advantage for use in racial comparison.

(e) To be fully qualified, the criterion should not be subject to a high rate of mutation. No investigations have been made on mutation rate in dermatoglyphics, but the character of dermatoglyphic variation suggests that mutation is a negligible factor in the production of racial differences."

Our interest in the criteria of classification goes beyond the merely classificatory. We want to know not only what kinds of men there are biologically, but also, if we can, in what order these kinds were evolved. As Morant puts it (1939, p. 72): "The main aim of physical anthropology is to unravel the course of human evolution, and it may be taken for granted to-day that the proper study of the natural history of man is concerned essentially with the mode and path of his descent."

Assuming that dermatoglyphics are valid criteria for the classification of man, can we make them perform this additional task of showing us "the mode and path of his descent"? Before such a task can be performed, the phylogeny of the dermatoglyphics themselves must be known. This necessarily subsumes a standard by which to

judge the relative primitiveness [2] of the various configurations, one that does not involve assumptions about the populations in which these configurations predominate.

In the discussion that follows, the only dermatoglyphics to be considered will be the digital patterns, to the exclusion of palmar and plantar configurations.

PHYLOGENY

Embryological evidence

Until the present time the most important evidence that has been adduced to the phylogeny of digital patterns has been that from the field of embryology. Midlo and Cummins, in their work upon the primates (1942, pp. 161-168), state their belief that, of the three Galton patterns, whorls are the most, and arches the least, primitive, with loops intermediate. They arrive at this position from the following considerations: Digital patterns arise on foetal pads. Ordinarily, whorls are associated with the more elevated and more clearly defined pads. Since these are considered to be more primitive than lowered and less clearly defined pads, it would seem to follow that the patterns associated with them, the whorls, are also primitive.

But the authors concede that, for a number of reasons, "this position cannot be maintained without reservation." The most important of these reasons, for our present purpose, is at once apparent. We see that the interpretation of the phylogeny of the foetal pads is fundamental to our understanding of the phylogeny of the patterns, yet Midlo and Cummins themselves cast a doubt on their judgment upon the pads: "Whether the pads viewed as primitive are so actually," they say (p. 162), "is problematic. It is merely that these fetalized pads are of the form most widely present among mammals generally, and prevalence may not be a safe criterion of primitiveness." Finally, in the next sentence they seem to lose faith in the expression of phylogeny in human dermatoglyphics altogether: "Furthermore, the extreme differences in dermatoglyphic expressions among some human races illustrate that categories of primitiveness in the dermatoglyphics

[2] The terms "primitive" and "advanced" are to be taken in their chronological, non-invidious sense. To proceed otherwise would be to subscribe to what Richard M. Weaver, in his "Ideas Have Consequences" (*Chicago*, 1948) castigates as "the Whig theory of history," according to which the newer is necessarily the better.

of primates may be vitiated by the introduction of pattern-determining factors which are independent of those more directly related to a phylogenetic sequence." This problem of variability in man is of the greatest importance. We shall return to it in a later section.

We are left with the conclusion that the evidence from comparative embryology, while suggesting a certain sequence of development in digital patterns, is not yet beyond dispute.

Before taking leave of our discussion of the embryology of digital patterns, let us note that Bonnevie (1929) has shown that arches occur on a relatively thick embryonic epidermis, while loops and whorls occur on a relatively thin one. Is there any evidence that will tell us which of the two kinds of epidermis is the more primitive? Such evidence would be of great aid in interpreting the phylogeny of the volar pads as well as of the patterns.

Functional evidence

Dankmeijer (1938, pp. 380-381) notes that many writers have assumed that whorls are the most primitive patterns in man because patterns resembling them have been found in the palms of lower mammals, such as the prosimians and marsupials. "In my opinion," he states forthrightly, "this theory is wrong." The comparison here is between characters found in the palm, but not in the terminal phalanges of the fingers, of certain marsupials and prosimians, and characters in the terminal phalanges of man. The seat of the highly developed sense of touch has moved, in the higher primates and especially in man, from the palm to the digits; if the form of papillary patterns is associated with this sense of touch, then these too have migrated, in these higher mammals, from the palm to the digits, and therefore "certain papillary patterns of the digits in man cannot be considered primitive only because similar patterns are found in the palms of lower animals."

Midlo and Cummins make a similar observation (p. 187):

> "It has been shown that the uncertainties surrounding an inquiry into the affinities of man arise in part from the independence of evolutionary sequences involving different dermatoglyphic traits. . . . It is apparent that specializations of members and of regions are associated with responses in the dermatoglyphics. These responses may vary widely among genera within the same taxonomic group, as in great apes or New World monkeys. In consequence, it follows that the membral and regional specializations may so alter the dermatoglyphics that their traits come to be indicative of varying uses of members rather than or in addition to genetic affinities."

Thus the functional evidence, as interpreted in these two sources, may negate the evidence from embryology. Dankmeijer, however, having combatted one theory, is not prepared to suggest an alternative one. He feels that before the phylogeny of the papillary patterns can be understood, the functions of the different parts of the hand must be taken into account. "As long as complete insight into this question has not been gained, it is wrong to speak of more or less primitive patterns in man. In this connection the occurrence of the patterns in Pygmies has a remarkable aspect; they showed the lowest percentage of the so-called primitive form ever observed!"

Dankmeijer here touches upon the very point that aroused my own interest in this subject, for it seems to me that belief in the primitiveness of the whorl leads to anomalous interpretations with many other peoples besides the Pygmies. As to Dankmeijer's cautious refusal to propose an alternative hypothesis, I can best state my own belief by quoting from a personal communication from Cummins (1949): "His position is understandable," he writes of the passages quoted above, "though we cannot hope to accomplish productive thinking in such fields except through exploring by means of working hypotheses."

Geographic evidence

It is almost inevitable that we call our preconceptions to our aid in arriving at conclusions in matters of phylogeny. We approach the problem convinced that man is higher in the evolutionary scale than the other primates, the primates higher than all other mammals, and so on. So we find Midlo and Cummins saying (p. 163):

> **"In support of the validity of the chosen point of reference it should be recognized that if there be adopted any other sequence of patterns from the more primitive to the more specialized than that beginning with the whorl, an irreconcilable conflict with other standards of taxonomic grouping would result. The gibbon, great apes and man would then be assigned to a group more primitive than the monkeys and most prosimians."**

But we have already seen that, if we concentrate our attention on man himself, an acceptance of the suggested sequence leads to a conflict equally irreconcilable, with, as Dankmeijer exclaimed in the passage quoted, Pygmies showing the lowest incidence of the so-called primitive form ever observed. And even with the primates as a whole, things are not much more satisfactory, for, as Midlo and Cummins maintain in their closing words (p. 194), the dermatoglyphics in man are closer,

in certain fundamental respects, to those of the monkeys than to those of the great apes, and therefore "it is concluded that man stemmed from an ancestral stock more primitive than any recent ape, having dermatoglyphic traits more closely allied to those of the monkeys."[3]

The Zones and Strata Theory of Griffith Taylor

One difficulty with both the embryological and the functional methods of investigating the phylogeny of digital patterns may be that, with our attention concentrated at all times upon the organisms bearing these patterns, we are too close to our subject. I do not propose that we rid ourselves of our preconceptions altogether; what I now put before the reader is an entirely different set of preconceptions, a set known as the Zones and Strata Theory of Griffith Taylor. One immediate advantage of this theory is that the organism plays no direct part in it, though it must be admitted that there remain other, grave disadvantages inherent in the fact that it *is* a set of preconceptions.

For a full statement of the Zones and Strata Theory the reader is referred to the many books by Griffith Taylor in which it is expounded, notably his "Environment, Race, and Migration" (1937). Taylor has presented his theory, in briefer form, to biologists, in three papers in *Human Biology* (1930, 1936, 1941). In the present instance we shall content ourselves with the statement that, in its simplest terms, the Zones and Strata Theory proposes a centre of stimulus, or cradleland, at which are to be found the most recent manifestations of a given trait, and that surrounding this centre horizontally are zones of widening radius, in each of which the manifestations are successively more primitive, until in the outermost zone, or in regions that are for other reasons the least accessible to the centre of stimulus, the most primitive manifestations will be found. Archaeologically, traces of the most primitive condition may be found even in the centre of stimulus, but these will be overlaid by strata that are successively less primitive, until ultimately the surface is reached, where the most advanced condition

[3] When we place beside this statement the following one by Candela (1940, p. 217) selected from the extensive and still rapidly growing literature on the question of human affinities: "However, of all the living organisms thus far tested, only the anthropoid apes have been found to possess blood group factors which are identical with those of man. . . . It would seem to follow as a logical conclusion, therefore, that the blood groups of man and the great apes were inherited, together with the many other characters, from a common ancestor or ancestors," we see that the question is still open.

is found. There will be fewer and fewer strata as the outermost zone is approached, until finally only the most primitive stratum is left, and this is now on the surface. This theory, says Taylor, " is in effect the fundamental principle behind the 'age and area' theory of Willis in botany—or that propounded by Matthew in zoology and by Wissler in cultural anthropology " (1930, p. 35). Three applications are illustrated in Fig. 1.

This, then, is our preconception. All we need now do in our search for the phylogeny of digital patterns is to plot their distribution on a map, draw in as objectively as we can whatever isopleths present themselves, and make our interpretation solely on what we see in the map.

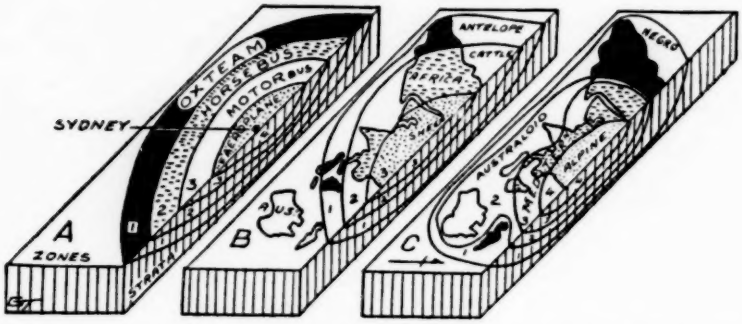

FIG. 1. THREE APPLICATIONS BY GRIFFITH TAYLOR OF HIS ZONES AND STRATA THEORY. (A) Application to the evolution of transportation in Australia, (B) of even-toed mammals (based on data by Matthew, 1915), (C) of the races of man (seen in larger scale and more detail in Fig. 5). (From Griffith Taylor, 1936, Fig. 1, p. 349)

The zones of pattern intensity

The *index of pattern intensity* (see Cummins and Steggerda, 1935, for the origin of this term) is a tool of the greatest use in mapping, for it is expressed in a single number, and sums up the condition on all ten fingers. This index gives to all whorls, no matter what their degree of complexity,[4] a value of 2, to loops a value of 1, and to arches a value of 0. Thus, the values of the index range from 20 to 0. If we accept the conclusion that whorls are the most primitive patterns,

[4] When, as in the present instance, data compiled by many different observers are to be used, the simplicity of the index is an advantage. I believe, however, that when greater uniformity of technique and interpretation is achieved, the index should be modified to discriminate between different kinds of whorls, and perhaps of loops as well.

then a high index, which indicates the presence of many whorls and few arches, must also be considered primitive, and a low index advanced.

In 1943 Cummins and Midlo gathered substantially all of the then known data on digital patterns found in various peoples, calculated the indices of pattern intensity, and published the results as Fig. 145 in their textbook on dermatoglyphics (1943, p. 260). These results, with twelve additions from more recent literature, as noted, are listed in Table 1. For convenience in mapping and interpretation, the indices

TABLE 1

Indices of pattern intensity of sixty population samples, arbitrarily assigned to classes

(Adapted from Cummins and Midlo, 1943, Fig. 145, with twelve additions)

CLASS	INDEX	SAMPLE AND NUMBER OF SUBJECTS	INVESTIGATOR
1	9.87	Bushmen (South Africa); 32	M. Weninger, '36
(9.50 to 10.99)	10.18	Efé Pygmies (Belgian Congo); 350	Valsik, '38; Dankmeijer, '38
	10.94	Batwa (N. E. Kivu); 84	M. Weninger, '37
2	11.35	Hottentots (South Africa); 50	Fleischhacker, '34
(11.00 to 11.99)	11.41	Negroes (French Guinea)*; 317	Dankmeijer, '47
	11.53	Dutch; 2,980	Piebenga, '38; Dankmeijer, '34
	11.55	Quiocos (Angola)*; 107	Sarmento, '39
	11.63	Tibus (Cufra, Libya); 104	Sabatini, '35
	11.83	Norwegians; 24,518	Bonnevie
3	12.00	Negroes of Central Africa*	Valsik, '38
(12.00 to 12.49)	12.04	English; 7,000	Waite; Scotland Yard
	12.10	Flemish; 200	Piebenga, '38
	12.11	Sanga Negroes (French Equat. Africa)*; 234	Dankmeijer, '47
	12.12	Negroes (Portuguese colony); 275	de Pina, '36
	12.13	Danes; 101,511	Bugge, '32
	12.18	Negroes (Angola); 174	Sarmento, '41
	12.24	Russians; 22,000	Semenovsky, '27
	12.29	Portuguese; 5,000	de Pina, '36; Lopez; Valladares
	12.34	Nhembas (Angola)*; 113	Sarmento, '39
	12.37	Spanish; 10,000	Oloriz
	12.38	Negroes (Liberia); 401	Dankmeijer, '38; Cummins, '30

* Calculated from data in Dankmeijer, 1947.

TABLE 1 (*continued*)

CLASS	INDEX	SAMPLE AND NUMBER OF SUBJECTS	INVESTIGATOR
4	12.51	French; 15,000	Bayle, '26
(12.50 to	12.57	Maya[b]; 127	Cummins and Steggerda, '36
12.99)	12.62	Negroes (Libya)[a]; 250	Falco, '17-'18
	12.64	Germans; 300,000	Steiner, '36-'37; Karl, '34; Heindl, '21, '35
	12.73	Hungarians; 833	Bonnevie, '29
	12.79	Bakola Pygmies (French Equat. Africa)[a]; 233	Dankmeijer, '47
	12.85	Walloons; 200	Piebenga, '38
	12.88	Ainu; 55	Hasebe, '18
5	13.00	Chilean Indians; 246	Henckel, '34
(13.00 to	13.13	Javanese; 2,000	Dankmeijer, '38
13.49)	13.14	Italians; 2,129	Sabatini, '34; Falco, '08
	13.16	Batwa (Kivu); 174	M. Weninger, '37
	13.23	Negroes (East Africa); 1,300	de Zwaan
	13.45	Middle American Indians; 633	Cummins, Leche, Cummins and and Steggerda, '36, etc.
6	13.53	Rwala Arabs	
(13.50 to		(Jebel Shammar); 200	Shanklin and Cummins, '37
13.99)	13.69	Gilyaks (renamed Nivkhi) (lower Amur R.); 28	Suda
	13.70	Bayaka Pygmies (French Equat. Africa)[a]; —	Dankmeijer, '47
	13.84	Jews (New Orleans); 200	Cummins and Midlo, '27
	13.94	Battaks (Toba, N. Sumatra); 1,000	Maasland
	13.95	Eskimos (W. Greenland)[c]; 145	Cummins and Fabricius-Hansen, '46
	13.96	Mitwali (southern Lebanon); 138	Cummins and Shanklin, '37
7	14.03	Hindus; 50	Biswas, '36
(14.00 to	14.05	Eskimos (three samples: St. Lawrence Id., Point Barrow and W. Greenland); 273	Cummins, Hansen, Midlo and Cummins
14.49)	14.22	Six Nations Indians (Ontario)[d]; 41	Wolfe, '47
	14.33	Sumatrans; 500	de Zwaan, '08

[b] From Cummins and Steggerda, 1936.
[c] From Cummins and Fabricius-Hansen, 1946.
[d] From Wolfe, 1947.

TABLE 1 (*continued*)

CLASS	INDEX	SAMPLE AND NUMBER OF SUBJECTS	INVESTIGATOR
	14.34	Japanese; 10,208	Shute; Furuse; Kubo; Iseki; Hasebe; etc.
	14.46	Manchurian Chinese; 3,350	Furuse
	14.48	Koreans; 3,891	Kubo, '18; Takeya
8 (14.50 to 14.99)	14.57	North American Indians; 400	Cummins, Downey, Cummins and Goldstein
	14.81	South Melanesians (Loyalty Islands?); 35	Hesch, '32
	14.86	Formosan Chinese; 1,025	Kutsuna *et al.*; Matuyama *et al.*
	14.91	Chinese; 500	Kubo, '18
	14.94	Mongols (Barge); 138	Yokoh
9 (15.00 to 15.49)	15.08	Inner Mongolians; 191	Yokoh
	15.17	Buriats (Mongols) (Lake Baikal); 55	Yokoh
	15.39	Batwa (Kivu); 76	M. Weninger, '37
10 (over 17.00)	17.13	Eskimos of E. Greenland [c]; —	Abel, '33

have been arbitrarily grouped into classes.[5] The lowest indices are in Class 1 and the highest in Class 10.

While the number of population samples is far from adequate, and the designation of many of them very inexact, it is nonetheless instructive, and at the same time essential for our present purpose, to plot the distribution of the indices of the samples we have, on a map. This is done in Fig. 2.

[5] Class 1 has been given three times, and Class 2 twice, the range of the remaining classes. Further investigations may increase the number of samples with indices of pattern intensity at the lower end of the scale sufficiently to warrant breaking Classes 1 and 2 up into units comparable in range with the rest. There is a surprising gap between the two highest indices recorded—from 15.39 to 17.13. The latter, found in a sample of East Greenland Eskimos, is placed in Class 10 by itself.

FIG. 2. WORLD DISTRIBUTION OF INDEX OF PATTERN INTENSITY (BY CLASSES). The reader is invited to speculate on how he would draw the isopleths of the zones of pattern intensity on this map.

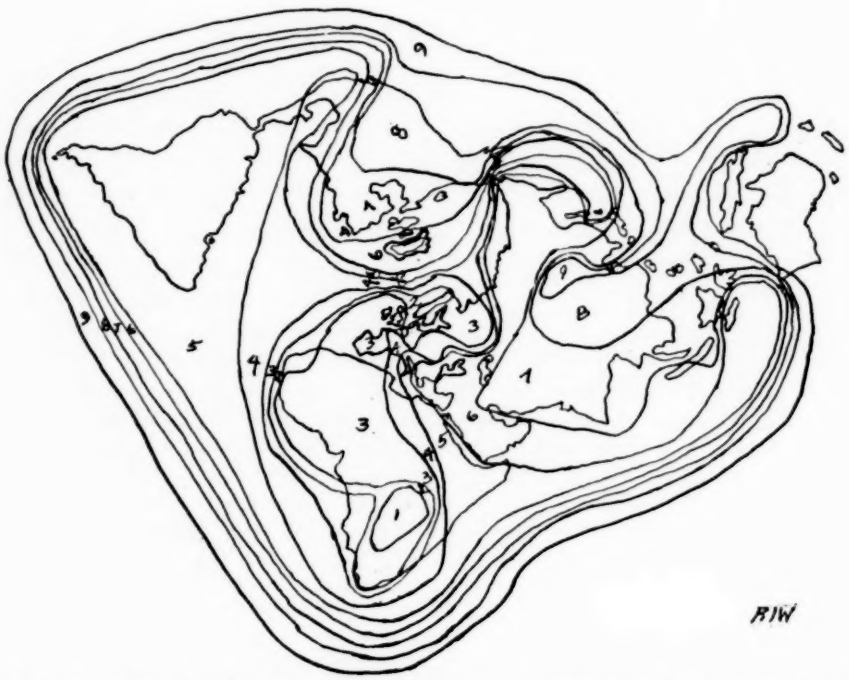

FIG. 3. ZONES OF PATTERN INTENSITY THAT RESULT WHEN ISOPLETHS ARE DRAWN AROUND THE LOCATION OF THE LOWEST CLASS OF INDICES AS CENTRE

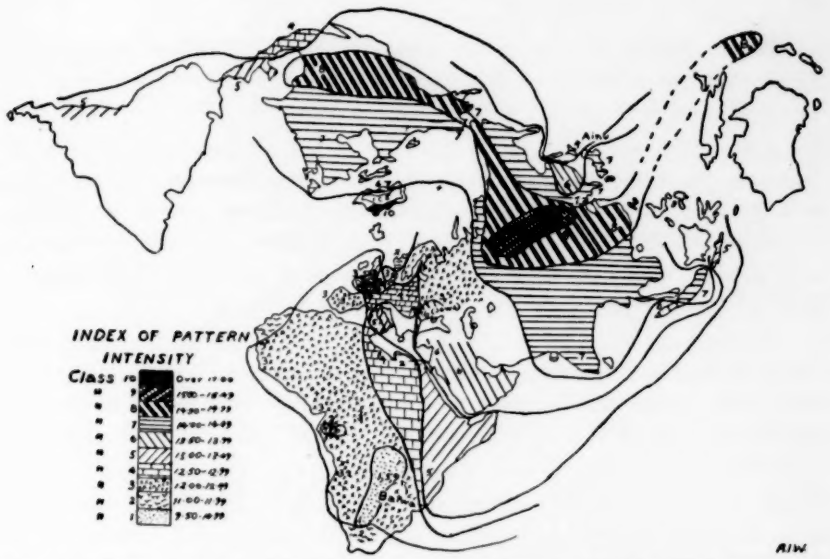

FIG. 4. ZONES OF PATTERN INTENSITY THAT RESULT WHEN ISOPLETHS ARE DRAWN AROUND THE LOCATION OF THE HIGHEST CLASS OF INDICES AS CENTRE. Class 9, the second-highest class, is assumed to be at the centre of stimulus, because the other classes group themselves around it in a logical sequence of zones. Class 10 contains but one index, that of a sample of East Greenland Eskimos; it is for the present disregarded.

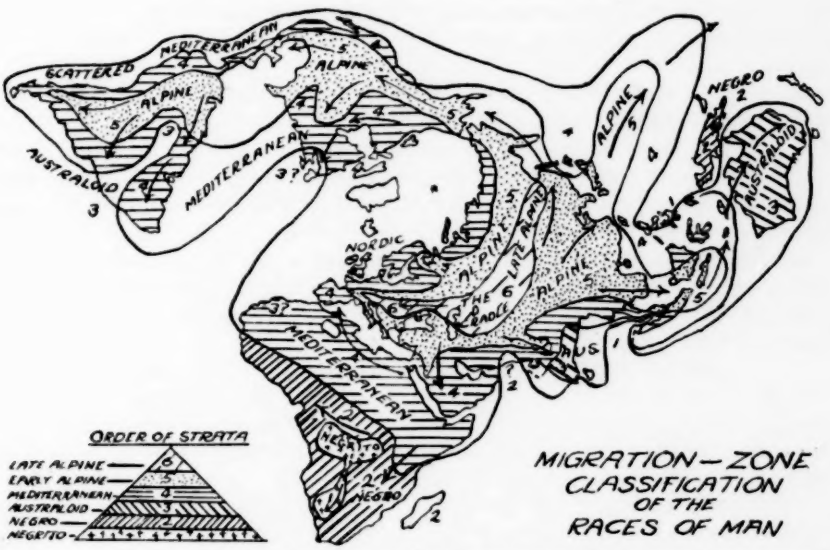

FIG. 5. COPY OF THE FRONTISPIECE OF "ENVIRONMENT, RACE, AND MIGRATION," BY GRIFFITH TAYLOR (1937). Titled, in part: "The logical classification of the Races of Man, based on racial history and ecological principles as well as on head index, hair, etc. The zones are generalized and indicate the primitive races in the region before modern migrations. . . ."

We see that superficially the map divides itself into two unrelated halves. In the main, the high intensities are found in Asia and North America, and the low in Europe and Africa. In order to apply the Zones and Strata Theory to the problem of the phylogeny of the digital patterns, we must find a relation between these two halves; we must look for a sequence of "zones of pattern intensity," as they may be called, such that either the lowest or the highest intensities are in the central zone, and successive classes of intensity are in successive zones, until at the periphery the opposite extreme is reached.

The best of many attempts to draw a map showing the zones that result from the assumption that the *lowest* classes of intensity lie in the central zone may be seen in Fig. 3 above. The result suggests a topologist's effort to diagram a solid rubber ball turned inside out, and offers no pattern that can be grasped by the eye. The pattern that results from the opposite assumption is, on the other hand, clear, as is seen in Fig. 4.[6] Our conclusion is, accordingly, that the central zone is occupied by the highest indices of pattern intensity. It follows, if the Zones and Strata Theory of Griffith Taylor is sound, that a high index of pattern intensity, and therefore a high incidence of whorls, is more advanced than a low index and a low incidence of whorls. Thus present assumptions as to the phylogeny of digital patterns are contra-indicated, and *the order of evolution is held to be from the primitive arches to the advanced whorls.* The assumption that loops are intermediate is retained.

Comments on some of the population samples

(1) Many samples are small and their designation imprecise, yet in some cases each is made to represent the greater part of a continent. On the basis of findings on fifty "Hindus," for example, we have placed the whole subcontinent of India and a good part of the rest of Asia in Zone 7.

---

[6] Many important resemblances between Figs. 4 and 5 will be seen. That construction of the former has been influenced by familiarity with the latter must be admitted. It is left to the reader to judge whether violence has been done to the facts thereby.

Obviously, the map in Fig. 4, for all the boldness of the isopleths and the definiteness of the zones, is tentative in the extreme; the conclusions drawn from it must be equally so. But to have been more diffident would not have guarded us against coming to the wrong conclusions; it would simply have made it more difficult to come to any conclusions at all. We may comfort ourselves with Cummins' statement, quoted above, that "we cannot hope to accomplish productive thinking in such fields except through exploring by means of working hypotheses."

But smallness of sample is in itself not objectionable, if it has been carefully designated both geographically and ethnically, and if the statistical treatment has been adequate; a too-large sample, insufficiently selective, such as the huge German samples, can yield results that are more misleading. Frequently, as with the samples from Italy, the authority cited has grouped together the results obtained on a number of population samples in one country, and recorded only a weighted average. Significant geographic differences within the country may have been masked in the process.

There are important areas in which no populations at all have been investigated, and which are accordingly left empty. Australia is one such area; South America, with the exception of Chile. is another; the area east of the Caspian Sea is a third. The lack of any dermatoglyphic data on the Australoids is unfortunate. Even more unfortunate is the gap near the Caspian Sea, for this, according to Griffith Taylor, is the cradleland of humanity.

(2) As already noted, the Eskimos of Eastern Greenland are the sole sample in Class 10. They have an index of pattern intensity of 17.13; this reflects the extraordinarily high incidence of whorls of 72.2 per cent, and makes these Eskimos far different from all other Eskimo samples tested. Cummins and Midlo state (1943, p. 259) that "Abel's material comprises several families, with as many as six, seven or eight siblings. It seems quite possible, therefore, that this series is not a representative racial sample." Cummins and Midlo accordingly omit it from racial comparisons, as we do in the present paper.

(3) The next highest index of pattern intensity is found in the surprising Batwa of Kivu (Belgian Congo)—surprising because the three groups of Batwa investigated exhibit in themselves almost the complete range of intensities in man, with one group near one end of the range, a second near the other, and the third exactly in the middle! A possible explanation for this extreme range, found in Pygmies generally, is discussed in a later section.

(4) The Jews of New Orleans have, in the present paper, been made representative of the population of south-central Europe. Cummins and Midlo, in their paper on these Jews, speak of their indeterminate migration history (1927, p. 91); we know, however, that most American Jews have come from central Europe. Siemens (1947, p. 79) concludes that Jews indigenous to various regions differ from each other in their main physical characters, and conform to a greater or lesser degree to the physical types of the populations that surround them; dermatoglyphic differences are also observed among the various Jewish groups (p. 82), "but it will remain for a future study to determine whether findings reflect the dermatoglyphic characters of the indigenous geographical stocks such as Russian, German, English, etc., from among whom the samples of Jewish prints were drawn."

That fundamental distinctions between the Russians and the Jews of Russia can be demonstrated is maintained by Manoiloff; in an arresting and insufficiently discussed paper (1927) he tells of numerous instances in

which he assigned individual samples of blood to their Russian or Jewish donors with an impressively high proportion of success.

After a careful analysis of the anthropology of central European Jews, Coon (1939, p. 646) concludes that, although they contain among them almost every racial type known to Europe, there is no doubt that the original Mediterranean blend of Palestine is the most important component in the Jewish constitution, accounting, at a guess, for more than half of the whole. Coon's view lends significance to both the difference between the intensities of the Jews and Russians and the similarity between those of the Jews on the one hand and the Rwala and Mitwali Bedouin, who live near the ancestral homeland of the Jews, on the other.

It may be safe to predict that the Jews of central Europe will regularly show higher pattern intensities than the surrounding populations. It would be interesting to investigate the dermatoglyphics of the Jews of China; if the above prediction is confirmed, can we further predict, in view of the high pattern intensities that prevail in the Far East, that these Sinicized Jews will have *lower* intensities than are found in their fellow Chinese?

(5) Cummins and Midlo remark on the close resemblance between the intensities of the Ainu of northern Japan, "a people of problematic relationship" (1943, p. 261), and the peoples of western Europe. This resemblance becomes significant when it is noted that in many other traits the Ainu are closer to the Europeans than to the Japanese. Here would seem to be a striking confirmation of the validity of dermatoglyphics as racial criteria.

(6) Among the most difficult samples to interpret is the Russian one. We shall leave comment on this and remaining samples until the next section.

RACIAL AFFINITIES

No attempt will be made to arrive at racial affinities purely on the evidence of the distribution of digital pattern intensities shown in Fig. 4. To make such an attempt would involve the danger of maintaining that the Danes are more closely related to the Negroes of Angola than to the Norwegians, or that the Italians are closer to the Javanese than to the Spaniards. This danger is recognized by Cummins and Midlo; they warn (1943, p. 264) that "it may not be justified to conclude that two races having equally high pattern intensity are on a par with regard to racial genealogy, or that a race presenting lower pattern intensity is farther removed in the genealogical tree." To eliminate this danger, we shall simply examine an existing theory of affinities, and judge the extent to which our new evidence supports it.

The Migration-Zones Classification of Races

Of the more than one hundred systems of racial classification that have been proposed since the first inclusive one of Bernier, in 1684, as listed by von Eickstedt (1937), there is one whose assumptions on racial affinities can most readily be tested by means of our new evidence from dermatoglyphics. This is the Migration-Zones classification of races, first proposed by Griffith Taylor in 1919. And the reason that it lends itself so readily to our purpose is that Griffith Taylor has used his Zones and Strata Theory, which is the tool we ourselves have used in interpreting the distribution of digital patterns, chiefly as a means of confirming his views on the order of evolution of the races of man, as expressed in his classification, and of tracing their migrations in prehistoric times. The zones in the resulting map can be compared with our own. Again the reader is referred to the books of Griffith Taylor or his papers in *Human Biology* for a detailed exposition of his findings.[7] The frontispiece of his "Environment, Race, and Migration" (1937) has been copied to provide us with the map in Fig. 5. In this map the "migration-zones" are highly generalized.

[7] The paper that appeared in *Human Biology* in 1941 is referred to by Cummins and Midlo (1943, pp. 265-266), and they give a valuable discussion of the implications of Taylor's findings for dermatoglyphics in the following words:

"Gross geographic trends of pattern intensity are recognizable. In Africa, Asia and the Americas there is a tendency toward progressive reduction of pattern intensity from north to south, while in Europe the line of reducing pattern intensity is from south to north.

"The possibility that ancient lines of dispersion might be reflected in dermatoglyphics is suggested by these distributions of finger-print types. Early human migrations have been traced by numerous investigators. Taylor . . . emphasizes that early migrations must be considered with the idea that the continents are three 'peninsular' offshoots from central Asia. The three peninsulas are Europe-Africa, southeast Asia with Australasia, and northeast Asia with the Americas. Significantly, the postulated center of distribution in Asia coincides with recent populations characterized by high pattern intensity. In passing northward in Europe, southward in Africa and southward in the Americas the populations exhibit progressively reduced pattern intensity. This reduction reaches its maximum in Africa, in the southern extremity (Bushmen) and in the central area (pygmies). Next in order of reduction of pattern intensity are populations of northern Europe, especially the Scandinavian countries. Though in the Americas there is a progression of diminishing pattern intensity from north to south, the level of intensity is relatively high; this is suggestive in the light of the supposition that migrations to the Americas were relatively late."

In the present paper the discussion is amplified.

There are, according to Taylor, five races of man, differentiated primarily on the basis of the cephalic index (whose value as a criterion for racial classification he defends against the attacks of Boas and others); these are, in the order of their evolution, the Negritoes, the Negroes, the Australoids, the Mediterraneans and Nordics (whom he groups together for certain purposes), and the Alpines, the most recently evolved of whom he calls the Late Alpines. He rejects the classic division of the peoples of Eurasia into the two races first named by Blumenbach in 1781—the Caucasians and the Mongolians. He calls it pre-scientific, and maintains that in fundamental characters the Mongols are similar to the Alpines, and should be classed as such.

The races of man have become distributed in the zones shown in Fig. 5 in the following manner: In Asia, just to the east of the Caspian Sea, there has been and remains a centre of stimulus, where the evolution of man, as of mammals generally, has had its most rapid tempo. Owing to environmental and other factors, chief among them the recurrent ice ages, there have been periodic migrations from this centre, and each newly evolved race has pushed its predecessors further toward the outlying areas. Thus we find the Negritoes, by consensus the most primitive of races, occupying Zone 1, which consists of inaccessible African wildernesses and isolated islands in the Pacific that may at one time have been joined to Asia by land bridges that have since disappeared. In Zone 2 are found the Negroes; these folk Taylor believes to be less advanced than the Australoids, who occupy Zone 3. The Mediterraneans and the Nordics he holds to be sufficiently similar to be treated as one race, and we find them in Zone 4. Finally, the Alpines occupy the two innermost zones, Zones 5 and 6, with the Late Alpines, the most recently evolved race of all, still occupying the cradleland itself. (The question of the relative accessibilities of the continents to the cradleland, important geographically, need not be discussed here.) It must be remembered that the migrations are held to have been essentially complete before the last ice age, and that at this time man was much more subject to ecological forces than he is today.

Let us now examine Figs. 4 and 5 in conjunction with each other. We shall try to determine the extent to which the evidence from the distribution of digital pattern intensities substantiates the migration-zone classification of races.

The most obvious similarity between the two maps is the orderly progression of zones in both. Zone 9, our innermost zone, is comparable in its geographic position, though not in extent, to Taylor's Zone 6.

Our Zones 8 and 7 are together similar to Taylor's Zone 5, though here there are some important differences that will be noted. Subsequent zones are less easily comparable. One reason is that, as the distance from the centre of stimulus increases, the populations that constitute the zones of pattern intensity become more widely dispersed. The small number of samples available to us will therefore hamper us more and more seriously as we approach the periphery.

Thus, our next two zones, Zones 6 and 5, are ill-defined, mainly because the number of samples in Classes 6 and 5 is small. (Later it will be shown that it is likely that the presence, in central Europe, of populations with indices in these two classes has been masked by the employment of too-broad sampling methods.) It is noteworthy, however, that the samples we do have are widely dispersed. They are to be found in each of the peninsular offshoots from central Asia—Europe-Africa, southeast Asia and the Americas—mentioned in Footnote 7 above. The agreement, in principle, with the migration-zone classification is manifest.

Classes 4, 3 and 2 are well represented in only one of the three peninsulas, Europe-Africa, with the result that only segments of the zones they occupy can be shown. It may be, however, that as more populations are investigated, these zones will be completed by the inclusion of parts of Australasia and perhaps of South America in them.

Zone 1 occupies exactly the same position in Africa in both maps. Since the Negritoes, who so far have been the only race in which Class 1 intensities have been found, have been examined only in Africa, we do not as yet know whether the Negritoes in Asia will have similar intensities. As will be shown below, we have no way of foretelling what their intensities will be.

We shall now proceed to make more detailed comparisons between Figs. 4 and 5.

The races of Asia and North America

Our Zone 9 corresponds only to the eastern part of Taylor's Zone 6, owing perhaps to the lack of data about dermatoglyphics in central Asia. The important consideration is that, in both maps, there is but one centre of stimulus, and furthermore, that in both, this centre lies in Asia. It is possible that further investigation will show that the populations in the cradleland itself, just east of the Caspian Sea, have patterns of the highest intensity, perhaps high enough to justify us in creating a Zone 10.

We find that our Zones 8 and 7 are together in partial agreement with Taylor's Zone 5, but the differences are important. We see that the well-known extension of Asiatic characters into North America is confirmed in our map, though for lack of data we cannot say whether South America too is reached. The apparent similarity, in the two maps, of the extensions of the zones into the islands of the Pacific is spurious, for our extension of Zone 8 is into Melanesia, where the presumably Negro inhabitants of the Loyalty Islands have been found to have exceptionally high intensities (for Negroes), while Taylor's extension is into Polynesia, whose Paraeo-ean inhabitants he regards as Alpine.

The north-south stratification of eastern China into Zones 6, 7 and 8 roughly corresponds with the distribution of the three Chinese subraces, with the Huanghoïds to the north, the Changkiangids central and the Chukiangids to the south (Liu, 1937).

The extension of Taylor's Zones 5 and 6 into Europe as an "Alpine Wedge" is not paralleled by our Zones 7 and 8. This is the most important discrepancy between the two maps, and it may indicate that the distinctions between "Caucasian" and "Mongoloid" are more fundamental than Taylor believes. Should this be so, then the Alpine race will once more be restricted to Europe (with the exception of outlying populations in other parts of the world who resemble the European Alpines and may be classified with them) and another name must be found for the Asiatic races.

It is of interest to note that in Fig. 4 as in Fig. 5 northern Sumatra is in a lower zone than southern Sumatra.[8] In Java, however, there is not the same correspondence between the two maps. It is unfortunate that the greater part of the mainland of Asia is represented only by the fifty Hindus, and it might have been wiser to have left the whole area blank. It is likely, however, that further investigation will show that most of this area falls into Zones 6, 7 and 8. If Taylor is right in his belief that the pre-Dravidians of India are related to the Australoids, then much of India itself will fall outside these zones. In which zone it will be found cannot be predicted, for we have no Australoid sample to help us.

[8] The Battaks of Toba, in northern Sumatra, are characterized as a "Caucasic" people in the *Encyclopaedia Britannica*, 11th ed., 1910, article on "Indonesian."

European races

It is difficult to speak about the "races of Europe" in any case, and if one adheres to the definition of racial distinctions given in Footnote 1 above, it is impossible to do so. In the present instance, our European samples are designated as English, French, Walloon, German and so on. What races these samples represent it is impossible to determine. The more noteworthy is it, therefore, that when the zones of pattern intensity in Europe, as shown in Fig. 4, are examined, a familiar design does emerge. If we use the term "Alpine" in its conventional, non-Taylorian sense, then there does seem to be an "Alpine Wedge" extending into western Europe. We see that the peoples of northwestern Europe, in whom we may expect a large Nordic component—the Norwegians, Dutch, Danes, English and Flemish—have indices of pattern intensity that fall into Classes 2 and 3. In the Iberian Peninsula, both the Spanish and Portuguese, in whom the Mediterranean component will be strong, also have intensities in Class 3. (The Mediterranean area to the east provides us with samples that disrupt the design, but in western Europe it is intact.) Between north and south are the French, Walloon, German and Hungarian samples, which may be presumed to have a high proportion of Alpines in them, and their intensities are all in the higher Class 4. We may deduce that, had the latter samples been carefully selected, so that only those subjects in whom Alpine characters predominated were printed, then the indices of pattern intensity would have been higher than in the unselected samples—perhaps high enough to place them in Class 5 or 6.[9]

There are two European samples that disrupt the design so far

[9] That this deduction is probably sound is shown by the results of an exhaustive study by Poll, published in 1937, and cited by Cummins and Midlo (1943, p. 266). These results, state the latter authors, agree with the separation of the Germans into Nordics and Alpines; populations of the north and west have relatively many arches and few whorls, those in the south and east have few arches and many whorls, and those who are intermediate geographically are intermediate in their patterns as well. We are therefore safe in saying that, had the samples of Germans listed in our table been differentiated into these three groups, the Germans of the south-east would have had pattern intensities substantially higher than those found in the samples as a whole, perhaps as high as Class 5 or even Class 6. If this is the case, then the zones of pattern intensity 5 and 6 are more nearly complete than they appear to be in Fig. 4, and the most serious discrepancy between our map and that of Griffith Taylor is lessened, though not destroyed. The low intensity of the Russian sample then becomes an even greater anomaly than it now seems. (See next paragraph in text.)

constructed: these are the large Russian sample, with an inexplicably low intensity, and the Italian, in which the intensities are fairly high.

One would expect to see a much higher index of pattern intensity in the Russians than one does see (Class 3). Asiatic intensities are high, and one would expect to see the many invasions of European Russia by Asiatic peoples find expression in an index of pattern intensity appreciably higher than those in western Europe. Also, the discrepancy between the Russians, in Class 3, and the Jews who presumably were their neighbours, in Class 6, is too great, even though the high figure for the Jews reflects a problematical relation to the two Arab samples, both of which have intensities of the same class. Finally, what possible ecological explanation can be advanced for the striking anomaly that the low index for the Russians creates in Fig. 4? In spite of the large size of the sample, one is reluctant to accept the index given as final. There are sufficiently large populations of immigrants from eastern Europe in the United States and Canada to make a further attack on this problem possible.

The high intensity of the Italian samples (Class 5) may reflect a higher degree of pattern intensity in Mediterranean peoples than the Iberian samples would lead us to expect, in which case the comment of Cummins and Midlo (1943, p. 261) is of interest: "Unfortunately," they say, "the precise geographic origins of the individuals composing many reported samples are not known. However, it is evident that Nordics are distinguished by low pattern intensity and Mediterraneans by higher intensity, Alpines being probably intermediate." The authors may have been encouraged in this view by considering the relatively high indices (Class 6) of the Rwala Bedouin, nomads of the north Syrian desert, who are of pure Mediterranean race (Coon, p. 623), and of the Mitwali, a mountain people of southern Lebanon, "Mediterraneans of the same type found in Palestine and northern Arabia, brachycephalized through the agency of the Alpine race" (Coon, p. 624). But reference to Table 1 once more brings confusion; for here we see that the pure Mediterraneans, the Rwala, have a *lower* intensity than the Mitwali, who have acquired some Alpine characters. (The Jewish sample, it will be noted, is intermediate.) Finally, the Tibu of Cufra, Libya, a people with a mixture of Mediterranean and Negro characters, have an intensity as low as Class 2. With the Negritoes of high intensity and the Eskimos of East Greenland, the Tibu are the only folk we have not assigned to a zone.

Groups with affinities to the Europeans

The striking resemblance of the pattern intensity in the Ainu of northern Japan to that of Europeans, and its difference from that of the Japanese, have already been noted. Taylor believes that the Ainu are akin to the Nordics, and that they moved eastward from central Asia at the same time that the Nordics moved westward.

It will be noted (Fig. 5) that Taylor believes the peoples along the west coast of the Americas to be "scattered Mediterraneans." We have only three population samples in this area in Fig. 4, but all of them have appreciably lower intensities than are found in Asia or among the North American Indian and Eskimo samples tested, and in this they do approach the Mediterraneans of Europe.

The Australoids and the Negroes

One of the most unfortunate gaps in our knowledge of human dermatoglyphics is caused by our lack of data about the Australoids. A search of American literature uncovered no dermatoglyphic data about them. It is difficult to predict what class their pattern intensities will fall into. In accordance with the migration-zones classification, the Australoids should occupy the zone next to that containing the Europeans, and they should be followed by the Negroes. Yet we find that the Negro samples are mainly in Zone 3, thus having *higher* pattern intensities than the Norwegian and Dutch samples. Their total range is almost exactly similar to the European range. The sequence of zones of pattern intensity for African Negroes is entirely logical (if we disregard the Mediterranean-Negro Tibu of Cufra), with Zone 5 lying closest to the centre of stimulus across the Arabian Sea and Zone 2 lying farthest from it.

It may be that the Australoids will also be found in the range 2 to 5. In that case, another tool must be devised to differentiate between the digital patterns of the Europeans, Australoids and Negroes. Such a tool is suggested below.

The Negritoes

The last group to be discussed are the Pygmies or Negritoes of Africa. The distribution of their indices of pattern intensity is more readily seen in Table 1 and Fig. 2 than in Fig. 4. No relation between the Negrito indices and those of nearby Negro samples can be recognized.

While the Negroes in west and central Africa are uniformly represented in Class 3, the Negritoes have a baffling range of indices, a range comparable to that of all the rest of humanity.

The Pygmies have always offered difficulties to anthropologists. They cannot be related to the Negroes, because they differ from them in too many fundamental somatic characters, among them, as we see, the digital patterns. They have exceptionally high cephalic indices; yet, if we follow Dixon and relate them to the Alpines on this account (Taylor, 1936, p. 355), then we can bring few other somatic characters to the defence of our classification. It would, moreover, be impossible to explain such a relation on ecological grounds.

Griffith Taylor's solution to the problem of Negrito relations, arrived at without knowledge of the peculiarities in the dermatoglyphics of the Negritoes, is remarkable, in that it gives what would seem to be a satisfactory explanation of these peculiarities. His solution is simple: he boldly removes the Negritoes from the main stream of human development, and makes them a collateral line (1937, p. 273, Fig. 101). They are relics, he maintains, of a very ancient type of man, probably contemporaneous with Neanderthal man. The implications of this view for our present discussion are easily seen: the two human stocks have evolved independently, and in each the complete range of dermatoglyphics has been produced, though in the Neanderthal stock the earliest digital pattern intensities, those of Class 1, have not as yet been represented.

The Zones and Strata Theory will not help us to interpret the distribution of Negrito intensities. To the extent that evolution in some of the Negritoes seems to have been as rapid, in one character at least, as in the most recently evolved peoples, the theory breaks down. This need not trouble us unduly, unless the simple certainties of a monistic concept are dear to us. Besides, if we are dealing with two only distantly related stocks of men, then we cannot expect complete correspondence in the mechanisms governing their rates of evolution. This much may be guessed: The Negritoes were early pushed out from the centre of stimulus (assuming that they arose there); in the intervening millenia, most of them evolved very little, in dermatoglyphics as in other characters; a few showed some advance, and a very few showed great advance, in dermatoglyphics, until today the Negritoes exhibit as great a range as the rest of humanity. What ecological mechanisms were at work in this evolution cannot now be guessed.

Should further research establish the soundness of Taylor's view as to the separateness of the Negritoes from the rest of humanity—and, as we have seen, it may be that the evidence from dermatoglyphics helps to do so—then it may follow that they are not merely a race, but a whole family of races, which may be partially differentiated from each other by their dermatoglyphics. But the Negritoes remain a problem, and what we have said must remain conjecture until more evidence is gathered.

We have no data on the dermatoglyphics of the Negritoes of the Pacific islands. When the Semang, Aeta, Tapiro and Mincopi are examined, they also can be expected to exhibit a wide range of pattern intensities.

Conclusion

"The student of dermatoglyphics will not fall into the error of presuming that dermatoglyphic features now provide ready answers for refractory questions in the history of races, particularly when he sees that attacks from many quarters have failed to yield a completely knit scheme of the affinities of races." (Cummins and Midlo, 1943, p. 264.)

It is hoped that the discussion just concluded will have been found to provide a few "answers to refractory questions in the history of races"; we should not be surprised to find that it has posed some new ones. We may conclude that digital patterns, when their phylogeny is assumed to be as in the present paper, and when they are used in conjunction with Griffith Taylor's migration-zones classification of races, do offer valuable clues to racial affinities. There are suggestive similarities between the distributions of digital pattern intensities on the one hand, and of the races of man, as postulated by Griffith Taylor, on the other. The greatest disparities between the two distributions lie in Europe. These disparities may be real, or they may result from faulty techniques. If the latter is true, we may hope that use of improved techniques will lessen them. That they will entirely disappear cannot be expected.

AN INDEX OF PATTERN VARIABILITY

It may have occurred to the reader that, in the discussion of racial affinities, no use was made of the "strata" concept in Griffith Taylor's Zones and Strata Theory. It is true that the strata are an archaeological concept, and archaeological evidence on dermatoglyphics is necessarily

meagre. Yet we may be able to modify the concept, restate it in biological terms, and thus be able to fashion a tool suitable to our purpose.

We must bear in mind that there is a significant difference between the distribution of digital patterns in zones of pattern intensity and the distribution of races in migration-zones. The crucial point to be remembered about the latter is that the primitive races are now found precisely where they did *not* originate, and the cradleland of all races is now occupied *only* by those most recently evolved. All types of digital patterns, however, are found in all zones of pattern intensity, and only their frequencies differ from zone to zone. Thus the two distributions are not exactly comparable.

This very difference may be of advantage to us here. Turning to the archaeology of the human races, we find that the strata concept postulates that no remains of the more recent races will be found in the outermost migration-zones, while remains of the primitive races will be found in all zones, but at deeper and deeper levels, and overlaid by the remains of more and more recent races, as the cradleland is approached. We may restate the concept in the following words: The zones in which the most primitive races are found are vertically the simplest, while those in which the most recent are found are vertically the most complex.

We are now ready, with the help of the ideas formulated in the preceding two paragraphs, to restate the strata concept in biological terms. While digital patterns of all types are to be found in all zones of pattern intensity, they will be found in different proportions in different zones. Furthermore, the *intraracial variability* of the digital patterns will also differ from zone to zone, in a manner similar to that of the vertical complexity of the migration-zones. If we can devise an *index of pattern variability,* then future research may establish the truth of the following proposition:

> *The index of pattern variability varies directly as the index of pattern intensity and inversely as the ecological distance from the centre of stimulus.*[10]

A concrete example will clarify the meaning of this proposition:

[10] By "ecological distance" is meant the complex of factors, such as environmental changes and accessibility, that separates a race in its present location from the location where it emerged—the centre of stimulus, in most cases. The concept is reminiscent of the "social distance" used by sociologists in their discussions of the stratification of society.

The Bushmen have the lowest index of pattern intensity known; they occupy a migration-zone and a zone of pattern intensity remote from the centre of stimulus in Asia; we may therefore predict that the variability of their digital patterns will be low.

We have seen that the index of pattern intensity does not help us to differentiate between the Europeans and the Africans, or between the Negritoes and the rest of humanity. It may be that the index of pattern variability, if one is devised, will help us to differentiate between these groups.

NOTE: Probably the likeliest point of attack in devising the index is the whorl. But this brings up a problem: If research does establish the truth of our proposition, that the index of pattern variability varies directly as the index of pattern intensity, can this relation not be attributed, at least in part, to the fact that a high index of pattern intensity connotes the presence of more whorls, and therefore offers more opportunity for variability, than a low one? When the index of variability is devised this possibility must be taken into account.

A biological statement of the Zones and Strata Theory

Lest I be accused of proposing a monistic concept, let me disclaim any intention of excluding all but environmental stimuli toward racial evolution in man, or of maintaining that no other centres of stimulus can exist but the one in central Asia. The Migration-Zones concept of Taylor does take account only of environment, and does propose but one centre. Our own zones of pattern intensity also group themselves logically about the same centre, but this does not mean that, once the primary centre of stimulus has been left, a new region cannot become a secondary centre—perhaps for other than environmental reasons. (How else, for example, can the exceptional variability in the intensities of the Negritoes be accounted for?) A whirlpool may have countless eddies, and if we are too close to it only the eddies may be seen; nonetheless the whirlpool exists.

We know that mutation rates for individual traits are exceedingly variable. It may be, however, that differences in rates are magnified by the assumptions that the observer brings with him. We find, for example, that Shanklin and Cummins (1937, p. 363) are perplexed to see that, by their interpretation, North-American Indians are primitive in some dermatoglyphic features (among them, digital patterns) and advanced in others; while, in contrast, "the Rwala are more advanced in the very items wherein the Indians are primitive, while for the three

palmar patterns noted in the Indians as advanced, signified in lower frequencies of patterns, the Rwala are seemingly more primitive!" Cummins and Midlo note the same paradox in the hands of the Chinese (1943, p. 265). The same authors pay much attention to the functional "specializations of members and of regions [that] are associated with responses in dermatoglyphics" (1942, p. 187), just as Dankmeijer does. Shanklin and Cummins, in trying to explain the paradox, also advance the possibility of "a differential and independent behavior of various features," which presumably may or may not be a response to functional changes, but they make the important admission that the paradox "may be identified also *with a false presumption of primitiveness*" (p. 363, my italics). If the phylogeny of digital patterns proposed in the present paper is accepted, much of the difficulty disappears.

Investigation of other dermatoglyphic features along the same lines as those suggested here may remove the "uncertainties surrounding an inquiry into the affinities of man [that] arise in part from the independence of evolutionary sequences involving dermatoglyphic traits" (Midlo and Cummins, 1942, p. 187), by proving that the evolutionary sequences are not so independent after all.

Finally, Shanklin and Cummins maintain that dermatoglyphic traits vary independently of both blood groups and cephalic indices (1937, p. 364). It is in a discussion of this statement that we shall come to the heart of the problem. We must agree that, if individual comparisons are made, then it is unlikely that specific rates and directions of variation will be matched in all three characters. Yet we have demonstrated, however tentatively, important resemblances between the distribution of digital pattern intensities and the migration-zones of races of Griffith Taylor. Since the cephalic index is Taylor's most important racial criterion, the migration-zones map may truthfully be called a "map showing the distribution of cephalic indices in man," and therefore we may claim that our highly generalized maps (Figs. 4 and 5) suggest—though they do not prove, because of insufficient data—*a fundamental relation between the distributions of digital patterns and the cephalic index.*[11] The same is true of the relation between the

[11] Newman, Freeman and Holzinger (1937, p. 87) find that similarities in finger prints are much greater between identical twins than they are between fraternal twins; in head widths, however, there is not much more similarity between the identical twins than between the fraternal twins, and head lengths are intermediate. This comparison suggests that digital patterns express genetic inheritance much more faithfully than does the cephalic index, and should there-

cephalic index and blood groups in Europe, as an analysis of Haldane's data (1941) that I hope to present on a later occasion, shows. And it is altogether likely that the same will be found to be true of the relation between blood groups and dermatoglyphics. It can be suggested, as an hypothesis, that any two racial criteria, if they are biologically valid and if they are compared on a sufficiently generalized base, will exhibit the same relation as the above. The operative word is *generalized.*

Any discussion about the mechanism at the base of this relation must lead us into the controversial question of environmentalism. That environmental stimulus is by no means the only factor influencing mutation rates is granted, and we must agree that "factors underlying racial divergence are complex, and it is impossible to assign shares to ancientness, degree of isolation and other factors which render the history of races so involved" (Cummins and Midlo, p. 264). Yet these concessions need not force us to abandon our thesis, that underlying all the complexities in the evolution of man introduced by the multitude of forces that have influenced it, lies a fundamental evolutionary pattern, a pattern created by man's ecological response to his environment.

This thesis has been stated by Griffith Taylor in geographical terms, as the Zones and Strata Theory. For biological purposes we may restate the theory in these words:

> *In man, the manifestations of racial criteria are more variable and phylogenetically more advanced as the ecological distance from the centre of stimulus decreases.*

Generalized, this may mean that:

> *The rate of mutation of any biological trait varies inversely as the ecological distance from the centre of stimulus.*

It must be left to future research to test the value of these hypotheses.

SUMMARY

1. Current assumptions about the phylogeny of digital patterns are reviewed. These have been made on the basis of inconclusive embryological evidence, and cannot be said to be definitive. Many difficulties arising from the assumption that the whorl is primitive are pointed out by Midlo and Cummins, by Dankmeijer, and by the present author.

fore be a better racial criterion. The putative relations between the distributions of digital patterns and cephalic indices become that much more noteworthy.

2. The digital pattern intensities of sixty population samples are plotted on a map of the world, and the distribution interpreted in the light of the Zones and Strata Theory of Griffith Taylor. It is tentatively assumed that the whorl is advanced and the arch primitive.
3. The light that the new assumptions about the phylogeny of digital patterns throw on racial affinities, as postulated in the Migration-Zones classification of Griffith Taylor, is shown. There is substantial agreement between the zones of pattern intensity and the migration-zones. In both, the central zone or centre of stimulus is in Asia. The greatest discrepancies occur in Europe, and these may in part be due to faulty techniques. The exceptional variability of Negrito intensities is possibly explained by Taylor's postulate that they are of a different stock from all the rest of humanity. Digital patterns are found to be of no value in differentiating between Europeans and Negroes.
4. It is suggested that an index of pattern variability be devised. It may be found that this index varies directly as the index of pattern intensity and inversely as the ecological distance from the centre of stimulus.
5. A biological paraphrase of the Zones and Strata Theory is proposed, viz. the rate of mutation of any biological trait varies inversely as the ecological distance from the centre of stimulus. More specifically, in man the manifestations of racial criteria are more variable and phylogenetically more advanced as the ecological distance from the centre of stimulus decreases.

BIBLIOGRAPHY

BONNEVIE, KRISTINE. 1929. Zur Mechanik der Papillärmusterbildung. I. Die Epidermis als formativer Faktor in der Entwicklung der Fingerbeeren und der Papillärmuster. *Zeitschr. Wiss. Biol. Abst. D. Roux' Arch Entwicklungsmech. Organ.* 117: 384-420. (From *Biol. Abst.* 5, No. 10, Oct., 1931, No. 25615).

BOYD, WILLIAM C. 1940. Critique of methods of classifying mankind. *Am. J. Phys. Anthrop.*, 27: 333-364.

CANDELA, P. B. 1940. New data on the serology of the anthropoid apes. *Am. J. Phys. Anthrop.*, 27: 209-222.

COON, CARLETON STEVENS. 1939. The Races of Europe. Macmillan (*New York*).

CUMMINS, HAROLD. 1949. Personal communication.

——— and VIBEKE FABRICIUS-HANSEN. 1946. Dermatoglyphics in Eskimos of West Greenland. *Am. J. Phys. Anthrop.*, 4, n. s.: 395-402.

——— and CHARLES MIDLO. 1927. Dermatoglyphics in Jews. *Am. J. Phys. Anthrop.*, 10: 91-113.

——— and ———. 1943. Finger Prints, Palms and Soles. Blakiston (*Philadelphia*).

——— and M. STEGGERDA. 1935. Finger prints in a Dutch family series. *Am. J. Phys. Anthrop.*, 20: 19-41.

——— and ———. 1936. Finger prints in Maya Indians. In CUMMINS, H., M. S. LANE, S. M. LECHE, R. MILDER, I. D. STEGGERDA and M. STEGGERDA. 1936. Measures of Men. Ten Specialized Studies in Physical Anthropology in Mexico, Central America, and the West Indies, pp. 103-126. Dep't. Mid. Amer. Research, Tulane University of Louisiana (*New Orleans*).

DANKMEIJER, J. 1938. Some anthropological data on finger prints. *Am. J. Phys. Anthrop.*, 23: 377-388.

———. 1947. Finger prints in African pygmies and Negroes. *Am. J. Phys. Anthrop.*, 5, n. s.: 453-484.

EICKSTEDT, EGON FREIHERR V. 1937. Geschichte der anthropologischen Namengebung und Klassifikation. *Zeitschr. für Rassenkunde*, 5: 209-282; 6: 39-96 and 151-210.

HALDANE, J. B. S. 1941. The blood group frequencies of European peoples and racial origins. *Human Biology*, 12: 457-480.

LIU, CHUNGSHEE H. 1937. A tentative classification of the races of China. *Zeitschr. für Rassenkunde*, 6: 129-150.

MANOILOFF, E. O. 1927. Discernment of human races by blood, particularly of Russians from Jews. *Am. J. Phys. Anthrop.*, 10: 11-21.

MATTHEW, W. D. 1915. Evolution and Climate. Ann. Acad. Science (*New York*).

MIDLO, CHARLES, and HAROLD CUMMINS. 1942. Palmar and Plantar Dermatoglyphics in Primates. *American Anatomical Memoirs, No. 20.* Wistar Institute of Anatomy and Biology (*Philadelphia*).

MORANT, G. M. 1939. The use of statistical methods in the investigation of problems of classification in anthropology. Part I. The general nature of the material and the form of intraracial distributions of metrical characters. *Biometrika*, 31: 72-98.

NEWMAN, HORATIO H., FRANK N. FREEMAN and KARL J. HOLZINGER. 1937. Twins—a Study of Heredity and Environment. University of Chicago Press (*Chicago*).

SHANKLIN, W. M. and HAROLD CUMMINS. 1937. Dermatoglyphics in Rwala Bedouins. *Human Biology*, 9: 357-365.

SIEMENS, GERHARD J. 1947. A study of certain genetic traits found in the dermatoglyphics of Jewish people. Unpublished Ph. D. thesis, University of Toronto, Toronto.

TAYLOR, GRIFFITH. 1930. Racial migration-zones and their significance. *Human Biology*, 2: 34-62.

TAYLOR, GRIFFITH. 1936. The zones and strata theory—a biological classification of races. *Human Biology*, 8: 348-367.

———. 1937. Environment, Race and Migration. University of Toronto Press (*Toronto*).

———. 1941. Races of the world—a discussion of recent classifications. *Human Biology*, 13: 390-397.

WOLFE, ROY I. 1947. Dermatoglyphics of the Six Nations Indians of southern Ontario. Unpublished M. A. thesis, University of Toronto, Toronto.

A GRAPHIC METHOD FOR COMPARING PERCENTAGES AND MEANS OF CONTROL AND EXPERIMENTAL GROUPS*

BY ROBERT SCHREK, M. D.

From the Tumor Research Unit, Veterans Administration, Hines, Illinois, and the Department of Pathology, Northwestern University Medical School, Chicago, Illinois

IN A study on the relationship of smoking and cancer, the percentages of smokers of several clinical groups were compared with the percentage of smokers for a control (1). To present these data and their statistical import to readers, many of whom were not statistically trained, a special graphic method was developed. As this method may be used for presentation of other statistical data involving experimental and control groups, the graph is described in detail in this paper.

Two types of control groups are used in clinical studies. Usually the control is distinct and mutually exclusive from all clinical groups. In the other type, the control consists of the summation of all the experimental or clinical groups. It is then necessary to compare the entire group with subgroups.

Comparison of the percentage of the control group and of subgroups. Let us first consider the simpler case where the control group includes all the individuals and the experimental groups are subgroups. An

* Sponsored by the Veterans Administration and published with the approval of the Chief Medical Director. The statements and conclusions published by the author are a result of his own study and do not necessarily reflect the opinion or policy of the Veterans Administration.

example of this is given in Fig. 1 which shows the percentage of men among patients with cutaneous squamous cell carcinoma of different sites (2). The abscissa of the graph represents the number of patients and the ordinates the percentage. A horizontal line shows the percentage of men (67 per cent) in the control group of 856 patients. The points

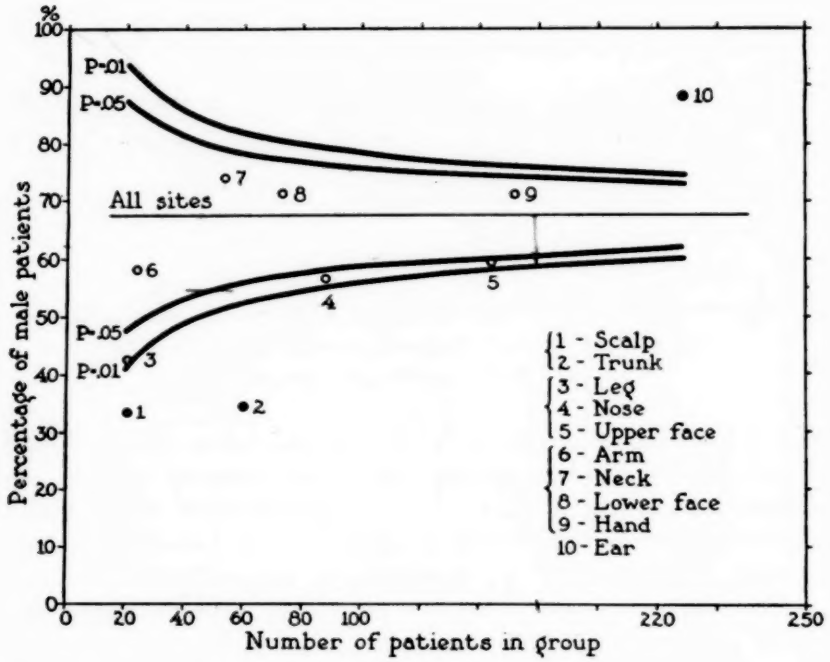

FIG. 1. AN EXAMPLE OF THE COMPARISON OF PERCENTAGES OF SUBGROUPS WITH THE PERCENTAGE FOR THE ENTIRE GROUP

The individual points in the graph represent the numbers and the percentages for each subgroup. The horizontal line indicates the percentage for the control or the entire group. The curves show the limits of significant percentages. Points outside the curves are significant.

in the graph represent the percentage and the number of patients for each site.

It is evident that the distance of the points to the control line indicates the difference between the percentage in the subgroup and the control group. The statistical significance of each difference can be determined by calculation or by a nomogram (3, 4). Percentages which

differ significantly from the control could then be indicated in the graph, as in Fig. 1, by means of closed circles as contrasted to open circles for non-significant percentages.

Another method of presenting statistical significance is to draw curves which will separate the points which are significant from those which are not. The curves can be calculated from the formula

$$p = P \pm \sqrt{\chi^2 \frac{N-n}{n} \cdot P\frac{(1-P)}{N}} \qquad (1)$$

where N = number of cases in control group,

P = percentage or fraction in control group,

n = arbitrarily assigned values for number of cases in subgroup,

χ^2 is determined from tables for a given value of statistical significance (for $P = .01$, $\chi^2 = 6.6349$; for $P = .05$, $\chi^2 = 3.8414$), and

p = calculated percentage corresponding to given n and χ^2.

Formula (1) is readily derived from the usual formula for χ^2 in the fourfold table with the symbolism:

Subgroup:	a	b	$n = a + b$
All others:	c	d	$N - n = c + d$
	$a + c$	$b + d$	N

$$P = (a + c)/N$$
$$p = a/n.$$

It can be seen in Fig. 1 that there are three groups (cancer of scalp, trunk and ear) with percentages which differ significantly from the percentage for all sites. Three other groups (leg, nose, and upper face) have percentages with borderline significance ($P = .01$ to $.05$).

The advantage of the graph is that all the groups can be readily compared as to (1) number of cases, (2) percentage, and (3) the significance of the difference of the percentage with the control percentage. One limitation of the graph is that it permits adequate comparison of the percentage of a group only with the control but not with all the other groups.

Comparison of percentages in control and in mutually exclusive experimental groups. In most cases the control group is distinct from

the experimental groups. An example is given in Fig. 2. The graph shows what percentage of skin tumors in different sites are basal cell carcinoma (2). The percentage of basal cell carcinoma for the control group (patients with cancer of upper face) is indicated by a horizontal line. As in Fig. 1, curves are used to indicate the limits of significance for percentages in the experimental groups.

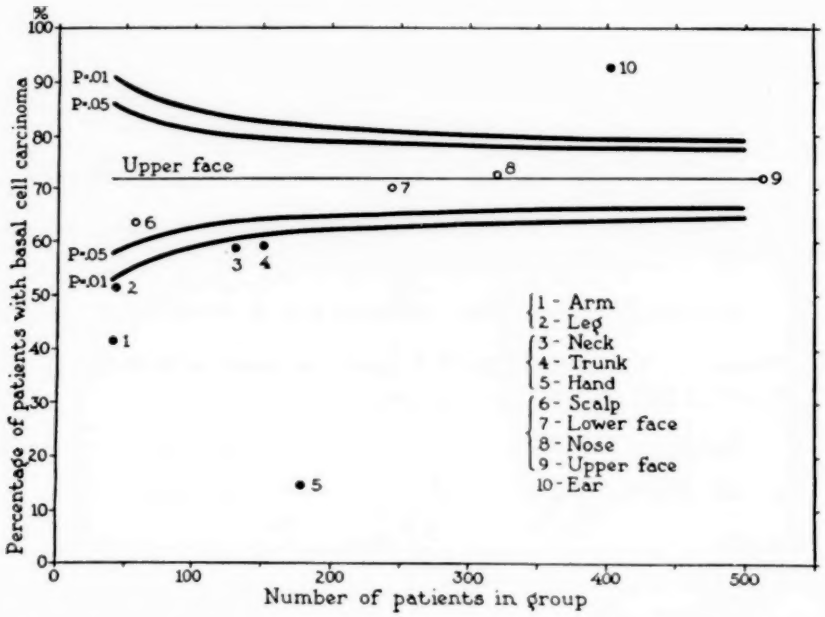

FIG. 2. AN EXAMPLE OF THE COMPARISON OF PERCENTAGES OF SEVERAL GROUPS WITH THE PERCENTAGE OF AN INDEPENDENT CONTROL GROUP

The calculations are somewhat more complicated than in the previous case. A reasonably simple computing scheme is the following:

Let the fourfold table be

Control	a	b	$n_1 = a + b;$	$p_1 = a/n_1$
Experimental	c	d	$n_2 = c + d;$	$p_2 = c/n_2$

$$P = (n_1 p_1 + n_2 p_2)/(n_1 + n_2).$$

We introduce the intermediate computing constant m, defined by

$$m = n_1(n_1 + n_2)/n_2\chi^2. \qquad (2)$$

Then for any desired n_2 we compute m and

$$P = \frac{2mp_1 + 1 + \sqrt{1 + 4mp_1q_1}}{2(m+1)}. \tag{3}$$

Finally, we get p_2 from

$$p_2 = \frac{(n_1 + n_2)P - n_1p_1}{n_2}. \tag{4}$$

These formulas may be readily derived if we write the formula for χ^2 as

$$\chi^2 = \frac{n_1(n_1 + n_2)(p_1 - P)^2}{n_2PQ}.$$

Comparison of the arithmetic means of control and subgroups. A similar method can be used to compare the arithmetic means of a control group and subgroups. The curves showing the limits of significance can be constructed from the following formula:

$$m_1 = M \pm ts\sqrt{(N - n_1)/Nn_1} \tag{5}$$

where

N = number of cases in control group,
M = arithmetic mean of control group,
s = standard deviation of control group,
t = normal deviate for desired significance (for P = .05, $t = 1.96$; for P = .01, $t = 2.58$), and
n_1, m_1 = number and mean for subgroup.

This expression may be derived from the usual formula for testing the significance of the difference between two means,

$$(m_1 - m_2)/s\sqrt{1/n_1 + 1/n_2} = t, \tag{6}$$

and the relations

$$N = n_1 + n_2,$$
$$NM = n_1m_1 + n_2m_2.$$

To represent the equation and the data graphically, let the abscissa indicate the number of cases and the ordinates, the mean. As previously, a horizontal line shows the mean of the control group. Points are used to indicate the mean and the number of cases in the experimental groups. Formula (5) is then used to construct curves showing the limits of sig-

nificance. Any point falling outside the curves indicates that the mean of the experimental group differs significantly from the control group.

Control and mutually exclusive groups. Usually the control group is separate and distinct from the experimental groups. In such a case computation may be done directly from the usual formula (6). It may be noted that this assumes that the control and experimental groups have the same standard deviations.

SUMMARY

A graphic method is proposed to facilitate the presentation of a comparison of experimental and control groups with respect to percentages or arithmetic means. The graph simplifies the presentation of the data and their significance and can be understood by a non-statistical reader.

REFERENCES

1. SCHREK, R., L. A. BAKER, G. P. BALLARD, and S. DOLGOFF. Tobacco smoking as an etiologic factor in disease. 1. Cancer. *Cancer Research,* 10: 49-58, 1950.

2. SCHREK, R. Cutaneous carcinoma. III. A statistical analysis with respect to site, sex, and pre-existing scars. *Arch. Path.*, 31: 434-448, 1941.

3. SCHREK, R. A nomogram for determining the statistical significance and the probable error of differences of percentages. *J. Lab. and Clin. Med.*, 25: 180-184, 1939.

4. ZUBIN, JOSEPH. Nomographs for determining the significance of the differences between the frequencies of events in two contrasted series or groups. *J. Amer. Stat. Ass.*, 34: 539-544, 1939.

NOTES

TEMPERATURE INFLUENCE OVER HUMAN GROWTH AND DEVELOPMENT

SURVEYING some 70,000 height, weight and menarchial records on American college students in 1937 (1) we encountered evidences of what then appeared to be the beginnings of a reversal in the human growth tide. This growth tide had brought about a 4-inch increase in stature of American men during the last two centuries and a steady, year-by-year, increase in average height and weight of freshmen matriculating in American colleges. College girls had exhibited the same stature improvement as well as a steady march of the menarche toward ever earlier years. The severe unseasonable warmth which prevailed widely from 1930 to 1936 seemed (in the 1937 survey) to have brought the growth tide to a halt or to a beginning reversal. Our 1949 rechecking of data from the same colleges now indicates that this reversal was definite but only temporary.

Studies on experimental animals under controlled temperature conditions had indicated that ease of body heat loss exercised a very evident dominance over growth and sexual development (2). We were thus much interested in knowing whether human beings in the mass were similarly dominated.

In Table 1 the data obtained in the 1949 survey have been added to those of the earlier period. It will be noted that the steady downward progression of menarchial age was reversed among University of Kansas girls born from 1920 to 1924, with a subsequent slow recovery to the early menarchial age of the 1919-born by those born in 1930. Similar changes occurred among University of Cincinnati girls, the regression in menarchial age in both places being roughly 0.4 years. This took place among girls approaching the menarche during the severely unseasonal warmth which prevailed at Cincinnati-Kansas latitudes from 1931 to 1936. Summers of 1934 and 1936 were characterized by especially

prolonged and severe heat, with few refreshing cyclonic interludes of cooler weather.

At the Women's College of the University of North Carolina the reversal in menarchial age trend was definite but less marked from the

TABLE 1

Changes in menarchial age of college girls

	NORTH CAROLINA		CINCINNATI		KANSAS	
YEAR OF BIRTH	Menarchial age	No. of cases	Menarchial age	No. of cases	Menarchial age	No. of cases
1921	13.42	320	..	..	13.20	13
1922	13.38	615			13.32	11
1923	13.40	379			13.54	24
1924	13.35	780			13.50	20
1925	13.28	839	13.75	24	13.44	31
1926	13.28	743	13.50	91	13.43	55
1927	13.23	681	13.04	272	13.36	93
1928	13.27	350	13.17	422	13.34	197
1929	13.37	83	13.15	386	13.17	231
1930	13.23	101	13.05	327	13.14	242
1931	..	..	13.27	231	13.16	61
1923-4-5			13.70	30	13.47	75
1890	14.16	347				
1891-95	14.15	1179	13.71*	194	13.69*	621
1896-00	14.05	1598				
1901-05	13.98	2321				
1906-10	13.85	3163				
1911-15	13.62	2892	13.52	1185	13.54	1938
1916-20	13.39	2648	13.23	1446	13.26	1362
1921-25	13.35	3665	13.18	295	13.37	98
1926-30	13.28	2014	13.16	1101	13.22	818

* 1890-1910

1919-born to the 1924-born girls. Summer temperatures there were less abnormal than those at Cincinnati and Kansas from 1931 to 1936. Mean July-August temperatures at Raleigh, N. C. were 1.5° above normal in 1934 and 1.7° above normal in 1936; corresponding elevations above normal at Cincinnati were 3.1° and 4.8°, and at Kansas City 8.4° and 10.1°.

Table 2 shows the remarkable gains in stature of American college freshmen occurring in the last 30 years. Roughly one inch has been added to the height of 17-year-old girls in that time and almost two inches to the height of boys. Weight gains have been 3 to 4 pounds

TABLE 2

Height-weight trends in college freshmen
(17-year-olds only)

YEAR OF ENTRANCE	KANSAS GIRLS			KANSAS BOYS			NORTH CAROLINA GIRLS		
	Height (in.)	No. of cases	Weight (lbs.)	Height (in.)	No. of cases	Weight (lbs.)	Height (in.)	No. of cases	Weight (lbs.)
1916-1920	..	..	..	68.45	..	135.60	63.61	404	116.38
1921-1925	..	..	..	68.32	..	135.61	63.62	1139	116.58
1926-1930	63.61	307	117.73	68.72	289	137.94	64.05	1396	117.99
1931-1935	64.37	354	118.92	69.19	545	141.60	64.23	1284	119.70
1936-1940	63.86	271	118.53	69.38	377	142.91	64.35	1762	119.81
1941-1945	..	..	..	..	..	..	64.79	1474	121.12
1946-1948	64.51	191	122.39	70.11	126	149.18	64.75	56	119.11

UNIVERSITY OF CINCINNATI FRESHMEN (17-18-19-year-olds)

MATRICULATING	Girls			Boys		
	Height	No. cases	Weight	Height	No. cases	Weight
1933-4-5	63.55	1024	121.68	69.17	1227	144.76
1948-9	64.47	617	124.12	70.23	824	152.80

TABLE 3

Height-weight recession and recovery in Kansas girls
(17-year-olds only)

YEAR OF ENTRANCE	HEIGHT	WEIGHT	YEAR OF ENTRANCE	HEIGHT	WEIGHT
1930	63.93	120.00	1935	64.73	116.76
1931	64.57	120.00	1936	63.15	116.54
1932	64.50	121.73	1937	63.85	117.08
1933	63.90	118.66	1938	64.59	121.96
1934	64.26	117.45			

18-year-old girls of 1934 class were 2.6 lbs. below 9-year average.
18-year-old girls of 1936 class were 3.1 lbs. below 9-year average.

for the girls and 9 to 14 pounds for the 17-year-old boys. Even in the midst of this general long-term improvement, however, the severe summer heat of 1934 and 1936 in Kansas brought a definite reversal among 17-year-old girls (Table 3). Those of the 1936 entering class were

over an inch shorter and 3.5 pounds lighter than those of the entering class of 1931. The 18-year-olds also showed a dip of 2.6 pounds in the 1934 entering class and 3.1 pounds in that of 1936.

DISCUSSION

It seems evident that human growth and development under natural climatic conditions exhibit the same sensitivity to severe heat-loss difficulties that experimental animals evince under controlled laboratory temperatures. Girls and animals both show delayed onset of the menses and retarded growth in prolonged heat. That girls should show greater heat suspectibility than boys is probably related to the lessened tolerance of women generally—perhaps because less physical activity and lower metabolic rate means for them a more sluggish vasomotor mechanism for bringing internal heat to skin surface.

Behind the widespread stature increase of American youth of all latitudes lie two factors of major importance. First is the general reduction in illnesses through the years of infancy and childhood, achieved by more scientific feeding practices and reduction of communicable diseases through immunization procedures. Second is the improved distribution of vegetables and fruits at all seasons and to all regions. The one major area remaining for still further improvement is that of acute respiratory infections, against which little headway has yet been made.

REFERENCES

1. MILLS, C. A. Further evidences of reversal in the human growth tide. *Human Biol.*, 13: 363, 1941.
2. OGLE, CORDELIA. Climatic influence on the growth of the male albino mouse. *Am. Jr. Physiol.*, 107: 635, 1934. *Also*: Adaptation of sexual activity to environmental stimulation. *Ibid.*, p. 628.

CLARENCE A. MILLS, M. D.
Professor of Experimental Medicine
University of Cincinnati